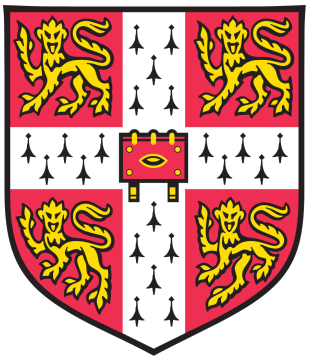
Investigating the role of ionised outflows in galactic evolution through spatially resolved spectroscopy

Robert Michael Gallagher



Cavendish Laboratory Astrophysics Group
Kavli Institute for Cosmology Cambridge
& Corpus Christi College
University of Cambridge

This thesis is submitted for the degree of Doctor of Philosophy
September 2019

Declaration

This thesis is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution. I further state that no substantial part of my dissertation has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution.

I hereby declare that this thesis is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the Preface and specified in the text. In particular, this dissertation is based on material published in:

- *Star formation inside a galactic outflow* Maiolino, R.; Russell, H. R.; Fabian, A. C.; Carniani, S.; **Gallagher, R.**; Cazzoli, S.; Arribas, S.; Belfiore, F.; Bellocchi, E.; Colina, L.; Cresci, G.; Ishibashi, W.; Marconi, A.; Mannucci, F.; Oliva, E.; Sturm, E. 2017, Nature, 544, 202.
- *Widespread star formation inside galactic outflows* **Gallagher, R.**; Maiolino, R.; Belfiore, F.; Drory, N.; Riffel, R.; Riffel, R. A., 2019, MNRAS, 485, 3409
- *The properties of ionised outflows and their host galaxies* **Gallagher, R.**; Maiolino, R. *in prep.*

Work based on these papers was conducted in collaboration with the listed co-authors.
This thesis does not exceed the word limit set by the relevant Degree Committee[a].

[a] As determined using TeXcount web service (version 3.1.1).

Investigating the role of ionised outflows in galactic evolution through spatially resolved spectroscopy

Robert Michael Gallagher

How feedback mechanisms trigger, regulate and suppress star formation in galaxies remains one of the main questions in astronomy today. Theoretical models and cosmological simulations of galaxy formation and evolution frequently employ feedback mechanisms from active galactic nuclei (AGN) and supernovae (SNe) to quench star formation in the early stages of massive galaxy formation and to explain the low efficiency of star formation in disk galaxies. A majority of this feedback is likely due outflows, ejecting large reservoirs of star forming gas from the galaxy and quenching star formation, powered by either AGN or SNe. However, recent theoretical and numerical models suggest that feedback due to outflows may also be able to promote star formation in gas in both galactic discs and entrained within outflows themselves.

This work focus upon studying the ionised component of outflows in the local Universe, with a particular focus on 'positive' feedback, where outflow feedback may trigger star formation in gas both in the galactic disk and inside the outflow itself, but also investigating the physical properties of galaxies and their negative feedback effect. To achieve this, I have developed a complex software specifically to explore multiple gaseous kinematic components in galaxies. I have applied this software to investigate outflows in the SDSS-IV MaNGA dataset, as well as data from other instruments.

In the first part of this thesis I discuss the first confirmed detection of star formation within a galactic outflow. By investigating the optical and near-IR emission line ratios of the outflowing gas, I unambiguously prove star formation is occurring within the gas. This is supported by the ionisation parameter of the gas, showing the gas excitation must come from UV radiation due to stars in the outflow and not stars in the galactic disk. Finally, I trace kinematical fingerprints of young stars blueshifted relative to those in the galactic disk.

The second part of this thesis focuses upon expanding the study of star formation in outflows in an unbiased sample of local galaxies. Through emission line ratios I find ~30% of our outflow sample show signs of *in-situ* star formation. This is supported again by the ionisation parameter, showing the gas excitation must be due UV radiation of young stars in the outflow. Investigating the relationship between star formation rates within the outflow and ionised outflow mass rates, I find a relationship implying large amounts of star formation should occur in high-z outflows. This new channel of star formation has numerous implications for galactic evolution.

The final part of this thesis focuses on general properties of ionised outflows and their effect on their host galaxies. I classify 168 outflows in local galaxies. Emission line ratios of the galactic disk show our outflow-host galaxies form a representative sample of local galaxies,

indicating outflow prevalence in all types of galaxies. I find a large fraction of the outflows are AGN-driven, though most are star formation-driven. Most star formation-driven outflows exist on or above the 'Main Sequence' in the SFR-M_* plane, suggesting the importance of starbursts for star formation-driven outflows. AGN-driven outflows exist on and below the main sequence, suggesting their efficiency at driving outflows in all types of galaxy. By studying the gas density through the [SII] doublet, I find outflowing gas is denser than gas in the galactic disk by a factor ~3, likely as a consequence of gas compression. Finally, by studying a number of properties of the outflowing gas, I find that some outflow properties (e.g. dynamical timescale) may be influenced by the properties of their host , whilst other properties (e.g outflow velocity) are more affected by the driving source of the outflow .

Contents

Acknowledgements

First and foremost, I would like to thank my PhD advisor and supervisor Roberto Maiolino. Throughout the duration of my PhD, Roberto's support, enthusiasm and generosity has known no bounds. He has always inspired me to follow my ideas in the pursuit of scientific knowledge, and has consistently challenged me with new opportunities. His expertise in the field has been paramount to my success in my projects.

As a part of Roberto's group, I have seen numerous people come and go, always bringing new ideas and beliefs to the group and often challenging my own perspectives. I have met some fantastic people throughout my PhD, and would like to thank Francesco Belfiore, Stefano Carniani, Michele Ginolfi, Andrin Fluetsch, James Trussler, Manuela Bischetti, Giacomo Venturi, Sim Brownson, Asa Bluck, Mirko Curti and Alice Concas (as well as many more people who have been a part of this massive group) for the fantastic discussions and lifelong friendships that I'm sure will last the test of time.

Additionally, throughout my time as a PhD student, I have met and collaborated with many interesting people, especially those who are part of the huge and ever growing MaNGA team. I would like to thank all of the numerous people from this group with whom I have had many long and deep discussions, often over a pint or two.

My time in Cambridge has been highly enjoyable and influenced by the numerous people I have met during my time in Corpus Christi College. I have made many lifelong friends, who have unequivocally provided me support throughout my PhD. They have always been there when called upon, and always ensured some of their time was put aside for us to relax and enjoy the many establishments on offer in Cambridge. Specifically, I would like to thank Izzy Beaudoin, Alex Sigston, Charlotte Elves, Leonard van Thiel, Tim Watson, Theo Björkmo who have been present in my life for the entire duration of my PhD and have helped keep me sane at times. There are many people here I have not mentioned, but have been pivotal in the memorable time I have spent in Cambridge.

My greatest thanks, however, have to go out to my family, who have supported me, both emotionally and financially, fully throughout my life and more recently my PhD. They have time and again been there when I needed their love and support, and their visits to Cambridge

have been some of my most memorable times here. I have explored parts of Cambridge I would have never otherwise visited with them, and they have shared in some of the most memorable experiences in my life. To my entire family, without your love and support, I would have never got this far in life. Thank you for everything you have done and continue to do for me.

The bulk of the work carried out in this thesis makes use of the SDSS-IV MaNGA survey data, and as such it would be remiss for me to not acknowledge the various institutes that have made the SDSS possible over the years:

Funding for SDSS-I-II and SDSS-IV has been provided by the Alfred P. Sloan Foundation and Participating Institutions. Additional funding for SDSS-II comes the National Science Foundation, the U.S. Department of Energy, the National Aeronautics and Space Administration, the Japanese Monbukagakusho, the Max Planck Society, and the Higher Education Funding Council for England. Additional funding towards SDSSIV has been provided by the U.S. Department of Energy Office of Science. SDSS-IV acknowledges support and resources from the Centre for High-Performance Computing at the University of Utah. The SDSS web site is www.sdss.org.

The participating Institution in SDSS-II include the American Museum of Natural History, Astrophysical Institute Potsdam, University of Basel, University of Cambridge, Case Western Reserve University, University of Chicago, Drexel University, Fermilab, the Institute for Advanced Study, the Japan Participation Group, Johns Hopkins University, the Joint Institute for Nuclear Astrophysics, the Kavli Institute for Particle Astrophysics and Cosmology, the Korean Scientist Group, the Chinese Academy of Sciences (LAMOST), Los Alamos National Laboratory, Max-Planck-Institut für Astronomie (MPIA Heidelberg), Max-Planck- Institut für Astrophysik (MPA Garching), New Mexico State University, Ohio State University, University of Pittsburgh, University of Portsmouth, Princeton University, the United States Naval Observatory, and the University of Washington.

SDSS-IV is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS Collaboration including the Brazilian Participation Group, the Carnegie Institution for Science, Carnegie Mellon University, the Chilean Participation Group, the French Participation Group, Harvard-Smithsonian Center for Astrophysics, Instituto de Astrofísica de Canarias, The Johns Hopkins University, Kavli Institute for the Physics and Mathematics of the Universe (IPMU) / University of Tokyo, Lawrence Berkeley National Laboratory, Leibniz Institut für Astrophysik Potsdam (AIP), Max-Planck-Institut für Astronomie (MPIA Heidelberg), Max-Planck-Institut für Astrophysik (MPA Garching), Max-Planck-Institut für Extraterrestrische Physik (MPE), National Astronomical Observatory of China, New Mexico

State University, New York University, University of Notre Dame, Observatário Nacional / MCTI, The Ohio State University, Pennsylvania State University, Shanghai Astronomical Observatory, United Kingdom Participation Group, Universidad Nacional Autónoma de México, University of Arizona, University of Colorado Boulder, University of Oxford, University of Portsmouth, University of Utah, University of Virginia, University ofWashington, University of Wisconsin, Vanderbilt University, and Yale University.

Finally, I would like to acknowledge the source of funding that has made the work in this thesis possible. I would like to thank the STFC for their financial support throughout my time at the University of Cambridge.

CHAPTER 1

THE ROLES OF FEEDBACK AND OUTFLOWS IN GALACTIC EVOLUTION

Between 1926 and 1936, Edwin Hubble developed one of the first widely accepted frameworks for the classification of galaxies - his eponymous 'tuning-fork' diagram (Hubble 1926b, 1936). Based entirely upon the morphological properties of galaxies, determined from high-quality photographic images taken on the Mount Wilson 100-inch telescope, the classification scheme is roughly split into two sequences: early-type elliptical galaxies and late-type spiral galaxies. Elliptical galaxies exhibit ellipsoidal shapes and smooth, nearly featureless images, assigned a number 0 through 7 by the classification scheme characterised by their ellipticity, and are supported entirely by the random motion of their stars. Conversely, spirals are composed of a central spheroidal bulge surrounded by a rotationally supported disk. Within the context of the tuning-fork diagram, spirals are divided into two sub-populations: normal and barred spirals, where barred spirals possess a bar of stars running through the bulge, with the spiral arms starting at the end of these bars. Further, the scheme assigns a letter *a* through *c* to spiral galaxies dependent on the compactness of their arms. Finally, where these two classifications merge, forming the 'tuning-fork' shape, we find lenticular (S0) galaxies. Hubble encountered issues classifying some galaxies within the context of his scheme, such as irregular galaxies and dwarf galaxies, and improvements to the classifications have been subsequently put forth.

Though self described by Hubble as "descriptive and entirely independent of any theory" (Hubble, 1926b), for a time the Hubble tuning-fork was thought to be an evolutionary sequence, with galaxies evolving from left to right (elliptical to spiral) across the diagram, hence

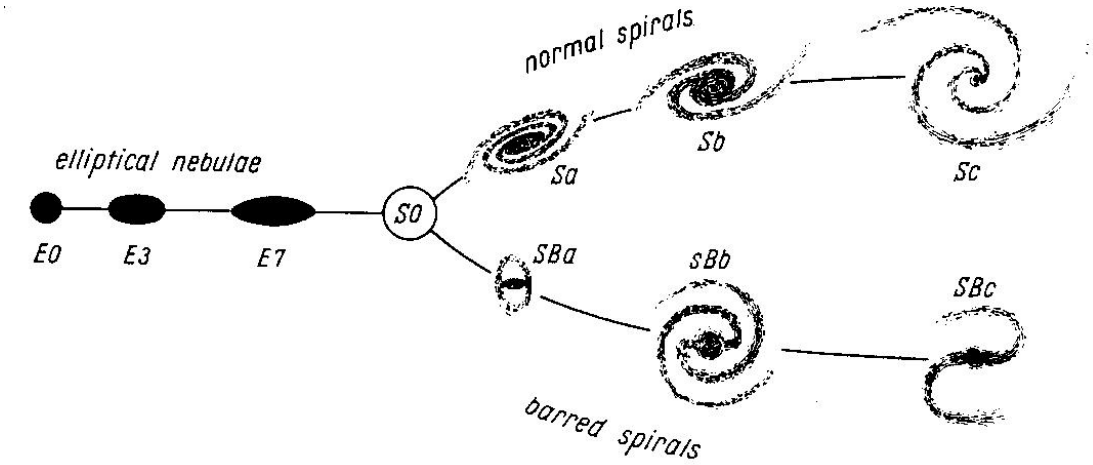


Figure 1.1: The Hubble tuning fork diagram. Starting from the left hand side, Elliptical galaxies are sub-divided into classifications from E0 to E7 dependent upon their observed ellipticity on the sky. At the intermediary point are S0, or lenticular, galaxies, wherein Hubble placed galaxies difficult to otherwise classify as elliptical or spiral, though nowadays typically intermediate bulge+disk galaxies are classified as such. Finally, on the right hand side, spiral galaxies are split into two distinct classifications: barred (SBa-SBc) or non-barred (Sa-Sc). As one progresses through these classifications, galaxies are observed to have smaller bulges with less clumpy and compact spiral arms. Figure from Hubble (1936).

the nomenclature of "early-type" and "late-type" galaxies being associated to their respective classifications. Though now known to be a gross oversimplification of the true evolutionary paths galaxies follow, the scheme lay foundations for many further studies, likely inspiring the discoveries of galaxy bimodality exhibited in many other properties, such as colour-magnitudes and SF rates (SFR).

More recently, results from some of the first Integral Field Spectroscopy (IFS) surveys have enabled classification of galaxies by also including kinematic information. A new global quantitative measure of morphology, termed λ_R, allows for the kinematic classification of galaxies in a robust way, insensitive to projection effects. This allowed for the splitting of galaxies into two distinct classes: fast- and slow-rotators (Emsellem et al. 2007; Cappellari et al. 2007). All fast-rotators are assumed to have a disk-like morphology, and if seen edge-on would appear morphologically with a range of spheroid fractions, ranging from disky isophotes (E(d)) to thin S0s. Conversely, slow-rotator early-type galaxies (ETG) are typically always found with ellipticities $\epsilon \geq 0.4$, with central cores in their surface brightness, and form a parallel sequence to sprial galaxies. As such, a revision of the tuning-fork diagram was proposed by Cappellari et al (2011b). Known as the 'comb' diagram, it combines elements of the tuning form diagram with the new findings enabled by IFS observations. As such, it is apparent how important kinematics are to the correct classification of galaxies, and how kinematics may

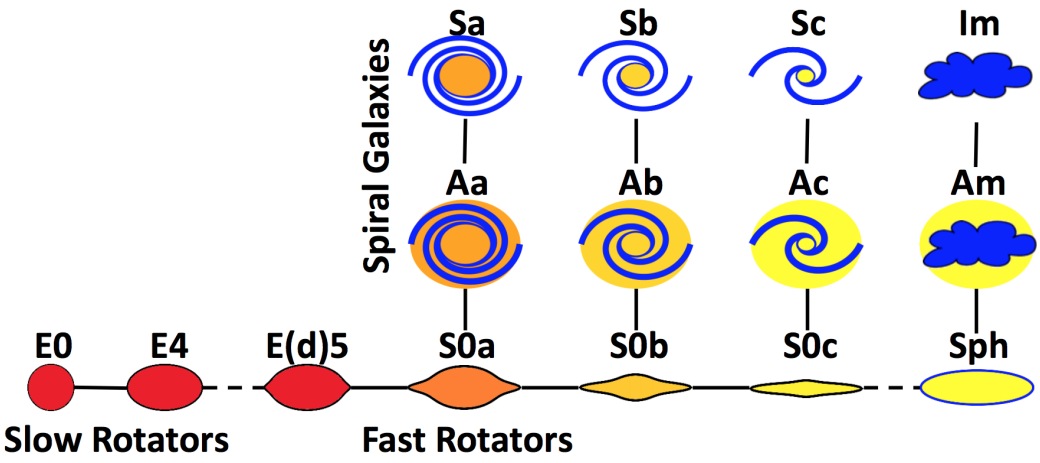


Figure 1.2: The "comb" morphology diagram of Cappellari et al (2011b), with the addition of Sph galaxies from Kormendy & Bender (2012). Fast rotators, if observed edge on, would appear morphologically within a range of spheroid fractions, from thin S0s, to flat ellipticals with disky isophotes E(d). They form a parallel sequence to spiral galaxies and few Sa galaxies have spheroids as large as those of E(d) gaalxies, indicating the bulge must grow in the transformation. Slow rotators form a "handle" to the diagram, and are always found to be rounder and have central cores to their surface brightness. The solid black lines are indicative of an empirical continuity, whilst the dashed line emphasises the dichotomy between fast and slow rotator ETGs. Figure from Cappellari (2016).

inform us of the evolutionary history of a galaxy.

This chapter provides an overview of the basic properties of galaxies, the role of feedback and outflows in driving the evolution of galaxies, and their associated observational probes and proofs, emphasising the key concepts upon which the work discussed in this thesis was built within the context of our current scientific understanding of the Universe. I begin with a brief introduction to the role of galactic feedback in our current Cosmological context, before moving on to discussing the effects feedback has on galaxies. We next begin to consider how to reconcile the theory of feedback by invoking feedback from both Star-Forming (SF) and AGN, with a particular emphasis on the extremes of AGN feedback in the context of suppression of SF. I next introduce the role outflows play in galactic evolution, discussing the multi-phase nature of outflows with a particular focus on ionised outflows, the phenomenon upon which the work in this thesis is based. Finally, I introduce the methods used to distinguish major sources of ionisation for both regions within galaxies and for outflowing gas, based upon their optical emission lines.

1.1 The formation of galaxies within a cosmological framework

The processes driving galaxies through their formation and subsequent evolution are underlaid by many complex mechanisms. Within a cosmological framework, such as the current standard model referred to as ΛCDM[a], these can be modelled as intricate interactions between physical processes that couple the three main energy density components of the Universe: baryons, dark matter and dark energy[b]. These components have shaped the geometrical curvature, mass and energy content, and distribution of structures of the Universe since the start of inflation, to what we observe today. Within this paradigm, small fluctuations in the otherwise homogenous dark matter distribution of the Universe grew with time, until they reached a sufficient size (partly determined by the cosmological parameters) to collapse and form gravitationally bound structures, known as dark matter halos, that were the progenitors of present day galaxies and galaxy clusters. Through the hierarchical accretion of matter and the merging of smaller structures, these halos grew. As the hot, accreted gas began to efficiently cool through radiative processes, cold, dense gas flowed inwards, reaching the conditions for formation of the first, massive and completely metal-free stars, and subsequently the forming of protogalaxies at the centres of the halos.

As cosmic time marches on, the galaxies that now inhabit the Universe evolve through a complex set of interactions pertaining to their baryonic matter. Primarily, these interactions involve gas, stars and active galactic nuclei (AGN, incredibly compact objects, whose extreme luminosity is powered by material accreting onto the super massive black holes (SMBH) located within the nucleus of a galaxy). In fact, feedback mechanisms due to both AGN and SF are considered crucial for explaining the many properties of observed galaxies. Thus gas, being the primary ingredient for both accretion and SF, is crucial in determining the evolutionary track of a galaxy. The intense amounts of energy that are released by BH accretion (AGN) and supernovae (SNe) can significantly affect the evolution of a galaxy by both the heating and removal of gas from the gas reservoir of the galaxy (outflows). In fact, outside of merging, most galaxies evolve through a quasi-steady equilibrium between SF, inflows and outflows (both AGN- and SF-driven) (Lilly et al. 2013).

Whilst one cannot determine the full extent that feedback processes have on the total

[a] Throughout this theses we assume a standard inflationary ΛCDM cosmology, as corroborated by results based upon Planck satellite observations (Planck Collaboration et al. (2018).

[b] In astronomical terms, baryons refers to any visible matter defined within the standard particle physics model. The origin and nature of both dark matter and dark energy, which together account for more than 95% of the energy density of the Universe, remains one of the open questions within astronomy today. Most of the mass component of the Universe is alluded to be dark matter, whose existence is supported by a wide range of evidence such as the dynamics of disk galaxies, whereas dark energy, sometimes referred to as vacuum energy, accounts for the perceived acceleration in the expansion of the Universe. This effect is analogous to that of the cosmological constant, Λ, in the ΛCDM cosmology paradigm.

evolutionary life of galaxy, a large number of properties exhibited by galaxies and observed phenomena cannot be fully delineated without invoking feedback in one form or another. As such, a marriage of theoretical frameworks, and semi-analytic and numerical simulations of galaxy formation and evolution with observational evidence obtained from multi-wavelength observational approaches is necessary. Throughout this thesis I use spectroscopic techniques to explore both the general properties of galactic ionised outflows, as well as to obtain some of the first evidence for an exciting new channel of SF: SF within galactic outflows themselves.

1.2 The role of feedback in galactic evolution

1.2.1 The observational evidence of feedback in action

1.2.1.1 The stellar mass function of galaxies

One of the first perceived discrepancies between observational measurements and theoretical values arises from estimating the stellar masses (or, equivalently, luminosities, due to the intrinsic scaling relationship between mass and luminosity) of galaxies. Within the ΛCDM paradigm, the properties of a galaxy should evolve coincidentally with those of the dark matter halo within which they formed over cosmic times. However, if one computes the observed stellar mass function, by finding the number density of galaxies with certain stellar masses (or luminosities), and compares this with the dark matter halo mass function predicted by ΛCDM models under the assumption that baryons distribute proportionally to the mass within dark matter halos, disparities arise at both the low and high mass ends (Silk & Mamon, 2012; Bower, Benson, & Crain, 2012; Papastergis et al. 2012; Kormendy & Ho, 2013). The observed halo mass function is best fit by a Schechter function (Schechter, 1976) of the form:

$$\phi(M) = \phi_* \exp\left(-\frac{M}{M_*}\right)\left(\frac{M}{M_*}\right)^{\alpha} \tag{1.1}$$

where $\phi(M)$ is the number of galaxies per unit mass, M_* is the characteristic mass at the inflection point of the function, after which the function changes from a simple power law to an exponential, and α is the slope at the low mass end. In effect, the ΛCDM model predicts a larger number of low-mass and high-mass galaxies should be observed than there are. Additionally, discrepancies at all masses suggest more baryons should be contained within galaxies than we detect. This comparison can be seen in figure 1.3.

Some amount of these shortfalls are believed to be either baryons contained within a Warm-Hot Intergalactic Medium (WHIM, Davé et al. 2001) or in cooler gas not yet accreted onto galaxies. However, to fully explain these discrepancies, generally models invoke different modes of feedback at different galactic stellar masses. The minimum of the discrepancies falls around

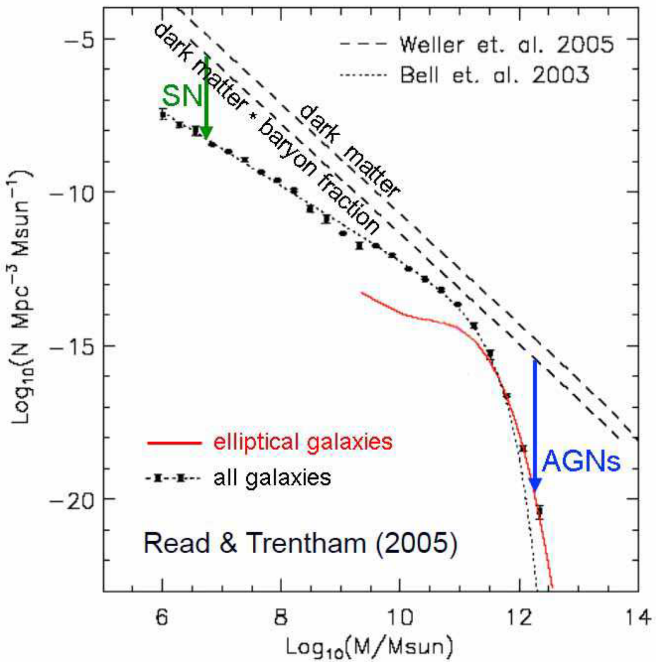
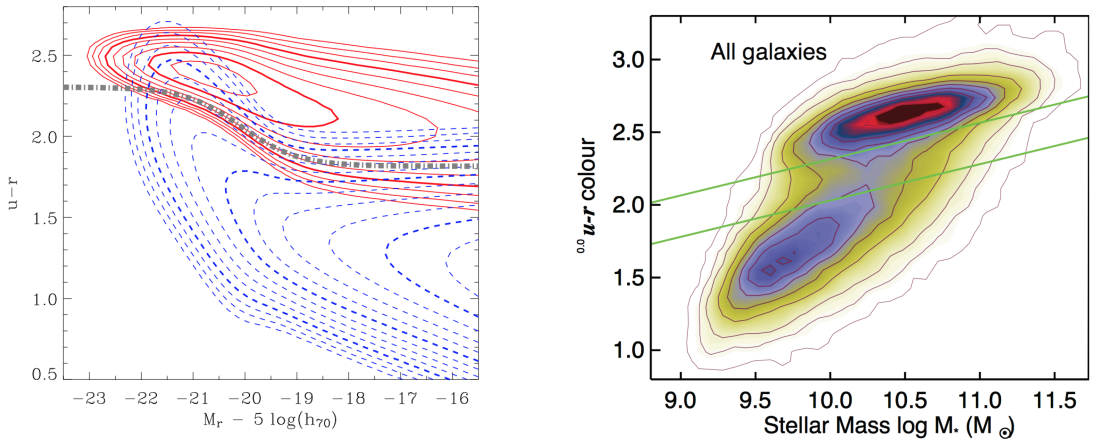


Figure 1.3: The observed galaxy baryonic mass function, accounting for both stars and gas, of nearby galaxies, as determined by Read & Trentham (2005). Black points represent all galaxies, and are fitted by a Schechter function (dashed black line), whilst the red curve shows the function for elliptical galaxies alone. In comparison, the mass function of simulated CDM halos (Weller et al. 2005), as well as this multiplied by the cosmic baryon fraction ($f_b = \Omega_b/\Omega_m = 0.163$, Hinshaw et al. 2013). The green arrow indicates the effect of stellar (SN) feedback at the low-mass end, whilst the blue arrow indicates the effect of AGN feedback at the high-mass end. Without these feedback modes, semi-analytic models fail to reproduce the observed discrepancies. Figure from Kormendy & Ho (2013).

$3 \times 10^{10} M_\odot$. Moving towards smaller galaxies (lower masses), the progressively increasing lack of baryons is believed to be an effect of stellar feedback, typically through the heating and ejection of gas by SNe-driven winds, suppressing SF by removing the gas reservoirs. However, as we progress to higher masses and past the turnover mass, SNe feedback is inefficient and cannot explain the exponential growth of the observed baryonic mass discrepancy. At these higher masses, feedback from AGN (and gravitational shocking of the cosmological gas infall) serve to suppress SF through the effects of heating, relativistic jets and outflows, trapping baryons in hot gas or simply removing baryons from the galaxies. If these feedback mechanisms at all mass scales are taken into account, one can reconcile observations and models.

1.2.1.2 Observations of galaxy bimodality

The advent of large imaging and spectroscopic surveys within the last decade, such as the Sloan Digital Sky Survey (SDSS, York et al. 2000), have facilitated the study of the properties of



(a) (b)

Figure 1.4: (a) The bimodality of galaxies in the local Universe in the $u-r$ versus M_r diagram, shown as red and blue bivariate distributions. Figure from Baldry et al. (2004); (b) The distribution of galaxies in the $u-r$ versus stellar mass diagram for a local galaxy population, taken from the SDSS. The green lines delimit the green valley, located between the blue cloud (on the bottom of the diagram) from the red sequence (on the top). Figure from Schawinski et al. (2014).

galaxies derived from high quality, statistically significant datasets. Previous studies of the distributions of local galaxies through comparisons of their intrinsic properties, such as the $u-r$ colour as a function of the absolute magnitude, M_r, have highlighted the existence of two distinct populations within the diagrams, namely the "blue cloud" and the "red sequence" (Strateva et al. 2001, Blanton et al. 2003b, Baldry et al. 2004). These easily measurable quantities also happen to be roughly analogous to physical properties of the galaxies, with the absolute magnitude tracing stellar mass and the colour indices being affected by and thus tracing recent SF activity, namely:

- Young stellar populations due to large amounts of recent star formation (typically observed in low-mass, late-type spiral galaxies) contribute to blue colour indices
- Older stellar populations with no recent SF (typically observed in high-mass, early-type spheroidal galaxies) contribute to red colour indices

As such, the bimodality is further exemplified in the distribution plots comparing colour indices and stelar mass or absolute magnitude, as shown in figure 1.4.

Combining the information gleaned from these two diagrams, astronomers have suggested evolutionary paths that galaxies follow to produce this clear bivariate distribution. According to most scenarios, proto-galaxies formed at the centres of dark matter halos, characterised by enhanced early star formation, that dominate the low-mass, blue end of the distribution. As these

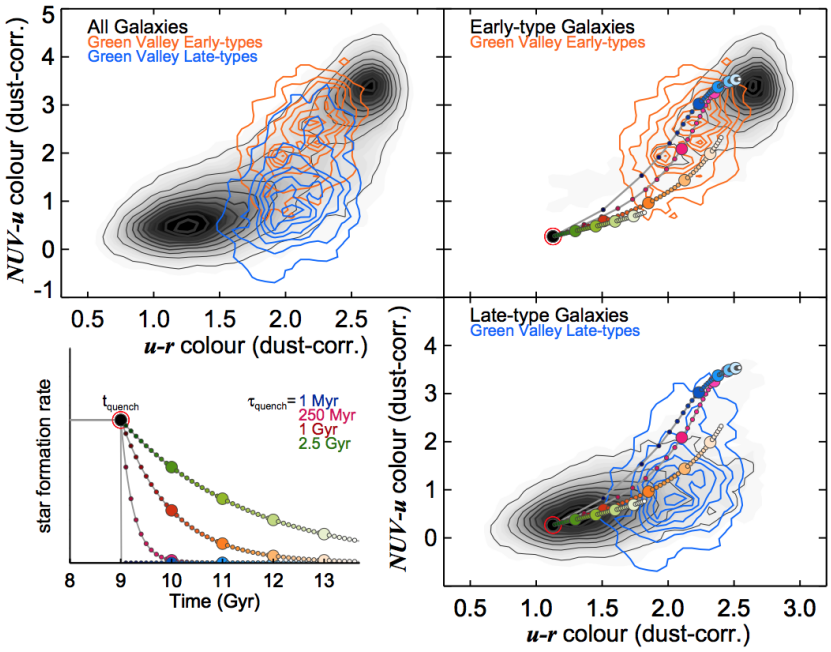


Figure 1.5: $NUV - u$ versus $u - r$ diagrams used to diagnose recent SF histories. These diagrams constrain the rate of change in the specific SFR, tracing quenching timescales. The right hand panels show the distributions of early-type and late-type galaxies in these diagrams, with their green valley counterparts shown as contours. Also shown are evolutionary tracks for galaxies with the same SF histories, until a time $t_{\rm quench}$ = 9 Gyr, at which point SF declines exponentially with quenching time scales $\tau_{\rm quench}$. The tracks show that different quenching timescales are required to reproduce the observed populations of early-type and later-type green valley galaxies. Figure from Schawinski et al. (2014).

galaxies evolve along SF sequences, they grow in mass both through hierarchical growth via mergers with other galaxies and through cold gas accretion from the IGM. However, if galaxies continued to evolve in this way alone, we would not see the distinct secondary population of red, quiescent galaxies spanning a range of masses, and we would also see a large population of massive galaxies with young stellar populations and ongoing SF in the local Universe.

As such, one or more processes must suppress the star formation within galaxies and prevent future bursts of star formation via either direct or indirect removal or heating of the gas reservoirs. This likely occurs once the galaxy reaches some threshold in stellar mass, explaining the lack of blue SF galaxies, but is also related to the environment within which galaxies live. To induce this quenching, a number of different mechanisms can be considered, such as starvation, whereby SF consumes the entire gas reservoir whilst further accretion of gas is denied (Peng et al. 2015), interactions the galaxy experiences due to the large scale environment within which it is embedded, such as ram pressure stripping (Fumagalli et al. 2014; Poggianti et al. 2016),

or, of interest within the scope of this thesis, feedback due to AGN activity (Cicone et al. 2014; Carniani et al. 2016), though SF-driven winds also serve to remove gas. Furthermore, this bimodality of galaxy populations in the colour-magnitude diagrams has also been shown to be prevalent out to $z \sim 1$, where colours are observed to be bluer and where the red sequence tends to be less populated, which serves to further confirm the migration of galaxies from the "blue cloud" to the "red sequence" from high redshift to the local Universe (Bell et al. 2004, Peng et al. 2010).

Upon deeper analysis of the diagrams, nestled between these two sequences exists another sparsely populated yet equally important area, colloquially referred to as the "green valley". Within this region, both early-type and late-type galaxies are found, with both exhibiting lower specific SFR (sSFR, SFR per unit stellar mass) than their blue, SF counterparts of the same mass, independent of morphology, suggesting they are in a quenching phase, transitioning from the blue cloud to the red sequence (Schawinksi et al. 2014). However, when one considers their $NUV - r$ colour, which traces recent SF as opposed to the more intermediate SF traced by the $u - r$ colour, early-types are found to have redder colours than late-types (figure 1.5). This evidence suggests two separate feedback mechanisms must quench galaxies, as early-types must be experiencing quenching on a rapid timescale (τ < 250 Myr), whereas late-types are experiencing a more gradual quenching timescale (τ > 1 Gyr) (Schawinksi et al. 2014). The likeliest explanation for these observations would be that starvation is pivotal in the evolution of late-type disk galaxies, yet AGN or starbursts provide the strong and fast feedback required to give rise to the red, quenched early-type passive galaxies.

1.2.1.3 The M_{BH}−host property scaling relations

Any theory of galaxy formation and evolution which is based within the current cosmological paradigm cannot ignore the contribution of AGN to the evolution of a galaxy's properties. In fact, the co-evolution of AGN and the bulges of their host galaxy has been suggested by tight correlations observed in local galaxies between SMBH masses and many bulge properties, such as the bulge mass (e.g. Merritt & Ferrarese, 2001; Kormendy & Gebhardt, 2001; Marconi & Hunt, 2003; Kormendy & Bender, 2009), bulge luminosity (e.g. Dressler, 1989; Kormendy, 1993; Magorrian et al. 1998; Kormendy & Gebhardt, 2001; Kormendy & Bender, 2011) and the velocity dispersion of the bulge (e.g. Ferrarese & Merritt, 2000; Merritt & Ferrarese, 2001; Kormendy & Gebhardt, 2001; Marconi & Hunt, 2003; Kormendy & Bender, 2011; Kormendy & Ho, 2013), commonly referred to as the $M_{BH} - \sigma_*$ relation, all of which are theorised to be regulated predominantly by AGN feedback (figure 1.6). The tightness of these relationships over a range of orders of magnitudes indicate that the two must mutually exert an influence on each other. Furthermore, the close agreement between the mass function predicted from AGN

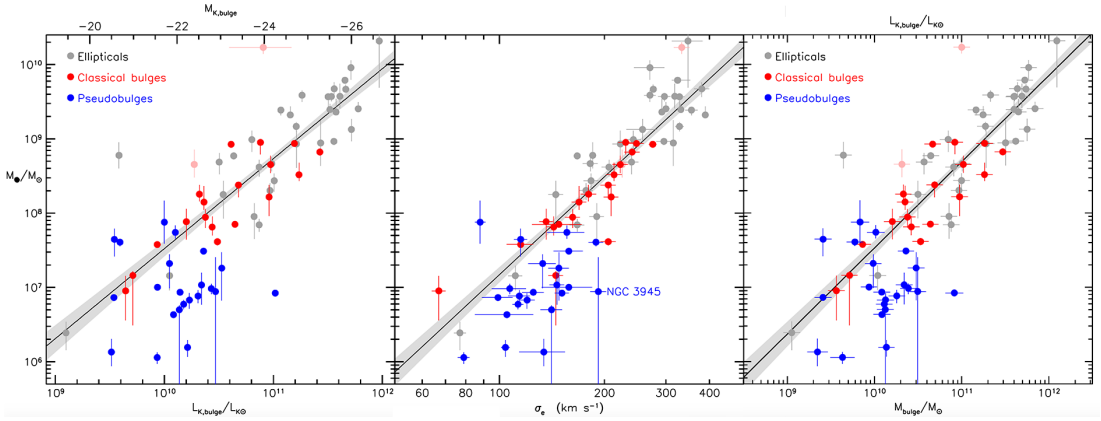


Figure 1.6: From left to right: BH mass versus host galaxy bulge K-band luminosity, versus stellar velocity dispersion and versus bulge mass. Shown as grey and red scatter points are elliptical galaxies and classical bulges, respectively, with the black line and shaded region representing the best fit and 1σ error, respectively. Also shown as blue points are galaxies with pseudobulges, which correlate only weakly with the trends observed for ellipticals galaxies and classical bulges. Figure from Kormendy & Ho (2013).

relics (Marconi et al. 2004) and that of the almost omnipresently detected SMBHs in local galaxies (Kormendy & Ho, 2013) suggests that any galaxy with a spheroidal component likely experienced a strong AGN phase at some point throughout their evolution (Mo et al. 2010).

Whilst for ellipticals the spheroidal component refers to the entire galaxy, for spirals this spheroidal component refers only to the central bulge, as evidence suggests that the SMBH masses correlate only weakly with "pseudo-bulges", and not at all with the properties of the disk.

1.2.2 Negative Feedback

Many of the effects of feedback mechanisms discussed previously have focused on the suppression and quenching of SF. This form of feedback is known as "negative feedback" to indicate the suppressive nature of the mechanisms with respect to the ongoing SF. Negative feedback can be associated with several mechanisms, driven by both star formation and AGN, such as galactic outflows, shocks ionising and heating the ISM and of the CGM. Isolating the relative role of each of these mechanisms in galaxy evolution is often difficult.

In the following sections, I provide a quick description of some of the primary negative feedback processes,

1.2.2.1 Stellar feedback: winds and supernovae

The processes at play throughout the life cycle of a star, from formation to death via SNe, can be a source of negative feedback upon their host. SNe inject copious amount of kinetic energy and momentum into their environment, heating or ejecting the cold gas in the disk. The quenching due to stellar feedback in low mass galaxies does reconcile the discrepancies between the observed galaxy mass function and that of the ΛCDM cosmology evolutionary framework at lower masses, but is insufficient to reproduce the observed lack of massive galaxies. This feedback mechanism is nowadays accounted for within numerical models and simulations of galaxy formation and evolution. The theoretical backbone upon which these simulations are built were derived from the ever growing wealth of high-quality, new data available, and consider the two most prominent modes of stellar feedback: the "momentum-driven" and "energy-driven" modes.

Within the "momentum-driven" mode, feedback mechanisms are dominated by the momentum imparted to the ISM by SNe explosions and UV-radiation young, massive stars. These momentum-driven processes are thought to be most effective at driving outflows of the dense, cold gas typical of dusty star-forming galaxies and AGN, as the transferred momentum cannot be simply radiated away in gas where radiative losses would be high. These feedback mechanisms, however, in order to be effective, require collisional coupling of the gas and dust within the outflows, as well as large optical depths to IR radiation not typical of normal, local Universe galaxies, and as such cannot be the sole feedback mechanisms responsible for the observed effects of stellar feedback. Conversely, the kinetic energy injected into the ISM via stellar winds from young OB and Wolf-Rayet stars and, much more importantly, SNe explosions in local galaxies is thought to be almost entirely dissipated radiatively, due to the low thermalisation efficiency (believed to be below ∼10%) of said energy (Chevalier, 1977; Murray et al. 2005; Thornton et al. 1998; Mac Low & Ferrara, 1999). In this scenario, powerful, large scale outflows are driven if the injected energy is sufficient to dissipate and unbind clumps of gas. This serves to self-regulate the SF within the discs of these galaxies (Chevalier, 1977; Chevalier & Clegg, 1985; Leitherer et al. 1992; Springel & Hernquist, 2003; Veilleux et al. 2005), and is of particular efficiency in low mass galaxies, where the gravitational potential well of the host dark matter halo is not as deep, thus gas within these outflows may possess sufficient energy to escape the galaxy itself, depleting the galaxy of vital cold SF gas.

1.2.2.2 Feedback from AGN

Whlist stellar feedback goes some way to explaining the discrepancies between the observed mass function of galaxies and that predicted by models in the ΛCDM paradigm, the feedback

mechanisms are unable to reconcile the observed lack of high-mass galaxies with the models. The required energy for SNe-driven winds to provide the necessary feedback is orders of magnitudes greater than that available from SNe explosions. Instead, AGN feedback mechanisms are introduced to explain the discrepancies at high masses, due to the copious amounts of energy they can exude to their surroundings. Following the energetics consideration laid forth in Fabian (2012), the energy released by the growth of a BH via accretion, assuming a 10% radiative efficiency for the accretion process, is calculated as $E_{BH} = 0.1c^2 M_{BH}$. Similarly, we can determine the binding energy of the dispersion supported galaxy bulge of the host galaxy, with a given mass of M_{gal} and a velocity dispersion of σ, from the virial theorem as $E_{gal} \approx M_{gal}\sigma^2$. Observations of BHs typically yield a BH mass, with respect to their host galaxy, of $M_{BH} \approx 1.4 \times 10^{-3} M_{gal}$ (Kormendy & Gebhardt, 2001; Merritt & Ferrarese, 2001; Häring & Rix, 2004), and most galaxies exhibit velocity dispersions $\sigma < 400$ km s^{-1}. Combining all of this information, we derive $E_{BH}/E_{gal} > 1.4 \times 10^{-4}(c/\sigma^2) \approx 80$. Thus, the evolution of any AGN host is likely to be significantly affected if even a small amount of the energy released by BH growth is transferred to the ISM[c]. This effect is nowadays included in many theoretical, numerical and semi-analytic simulations of galaxy growth and evolution (Kauffmann & Haehnelt, 2000; Granato et al. 2004; Di Matteo et al. 2005; Springel, Di Matteo & Hernquist, 2005; Bower et al. 2006; Croton et al. 2006; Ciotti et al. 2010; Hopkins et al. 2006, Scannapieco et al. 2012).

Similar to stellar feedback, AGN feedback is typically considered in one of two scenarios: "radiative mode" and "kinetic mode". The "kinetic mode" of AGN feedback serves to ensure that, once a galaxy has been emptied of gas capable of forming stars, the galaxy does not "refill" and subsequently ignite new SF, and that any gas either that either remains within the galaxy or is accreting onto the galaxy is kept hot and not allowed to cool. In many models, this is achieved by relativistic radio jets, associated with low accretion rates onto the BH, traversing the galaxy and injecting energy into both the IGM and CGM. This feedback is thought to be most important at later stages in the evolutionary cycle. The radio-jet kinetic feedback mode provides us with the most striking evidence of AGN feedback, manifesting as large bubbles in the x-ray emitting gas of the cores of clusters.

Whereas the "kinetic mode" is associated with low accretion rates, the "radiative mode" is associated with accretion that is radiatively efficient and close to the Eddington limit ($L_{AGN} \sim L_{Edd}$). This incredibly luminous phase releases large amount of energy, which can couple to the surrounding material and drive outflows with high mass outflow rates, quenching star formation within the vicinity by expelling the gas reservoirs required for continued SF. This

[c] Notably, this feedback only affects current and future SF, and does not significantly affect currently existing stars in the host.

feedback mechanism is thought to responsible for the observed $M_{BH} - \sigma_*$ relationship discussed previously. This mechanism was likely most prominent and effective at earlier times ($z \sim 1-3$, during the "cosmic noon" of the Universe) when galaxies contained large amounts of rich gas and quasar activity peaked, and likely acts on short timescales ($\tau \sim 10^7$ yr, Menci et al. 2008) of accretion.

1.2.2.2.1 Kinetic Mode

The kinetic mode of feedback provides us with the clearest observational evidence of the effect AGN have on their surroundings, such as bubbles and cavities in the x-ray emitting gas of the cores of some of the most massive galaxies, themselves typically the brightest cluster galaxy (BCG) in the centres of galaxy clusters. The hot gas surrounding these galaxies often has radiative cooling times that should ensure a cooling flow, with cooling rates of up to thousands of solar masses per year (Fabian 1994, 2012), which can be inferred from their X-ray luminosities via:

$$\dot{M} = \frac{2}{5}\frac{L\mu_m}{kT} \tag{1.2}$$

where L is the (mostly) X-ray luminosity of the gas, μ_m is the mean mass per gas particle at temperature T and k is the Boltzmann constant (Fabian, 2012). With no source to counteract this cooling, one would expect the gas to radiatively cool and collapse back upon the galaxy. Were this to happen, these galaxies would experience large starburst periods, appearing as bright, blue galaxies and growing in mass, hence populating the high end of the stellar mass function with young galaxies. Instead, these objects are observed as quiescent, red objects, and as such the AGN must be injecting energy directly into the surrounding gas to counteract the cooling flow (Peterson & Fabian, 2006; McNamara & Nulsen, 2007; Cattaneo et al. 2009). Through this process, theorists have reconciled the high mass end discrepancies of the mass function between observations and predictions from the ΛCDM model (Benson et al. 2003; Granato et al. 2004; Croton et al. 2006; Gabor et al. 2011, Bower et al. 2012).

The main mechanism through which the AGN inject energy into the surrounding medium, ionising and heating the medium and giving rise to the observed bubbles and cavities, are found to be intense radio jets propagating from the AGN (Boehringer et al. 1993; Carilli et al. 1994; Rizza et al. 2000; Bîrzan et al. 2004, 2012). Due to the low radiative efficiencies of these jets (typically between 10^{-2} and 10^{-4}), the jets mechanically inject their kinetic energy into the medium. The kinetic power delivered by the jets is typically in good agreement with the expected energy loss via X-ray radiation, thus demonstrating the energetics balance required for the feedback process to counterbalance the expected cooling (McNamara & Nulsen, 2007,

2012; Rafferty et al. 2006, 2008; Dunn & Fabian 2008, Fabian, 2012; Hlavacek-Larrondo et al. 2012, 2013). This mode is often also referred to as "maintenance mode" feedback, due to the extended timescales over which it acts.

1.2.2.2.2 Radiative Mode

Whereas the kinetic mode acts over long timescales to maintain a state of quiescence by preventing cooling of the ISM and CGM through the injection of kinetic energy, the radiative mode operates on short timescales during periods of intense BH accretion, releasing large amounts of AGN emission and sweeping material from the ISM in the form of powerful outflows. These large-scale outflows are generated by the hydrodynamical coupling of the ISM with the AGN radiation through sub-relativistic nuclear winds that develop from the innermost regions of the AGN accretion disk (Silk & Rees, 1998; King, 2003, 2010; Lapi et al. 2005; Menci et al. 2006, 2008; Zubovas & King, 2012, 2014; Faucher-Giguère & Quataert, 2012; Costa et al. 2014a, 2014b, 2015; Nayakshin, 2014). In this mode, AGN accrete close to the Eddington limit ($L_{\rm AGN} \sim L_{\rm Edd}$), generating winds that initially carry kinetic energies and momentum fluxes comparable to a few percent of the AGN bolometric luminosity (~0.05 $L_{\rm AGN}$) and to the radiative momentum output of the AGN ($\sim L_{\rm AGN}/c$), respectively. These highly supersonic ($v_{\rm wind} \sim 0.1c$) winds are destructive, shocking the ISM and driving gas out of the galaxy, depleting the galaxy's SF reservoirs of gas.

These powerful outflows have been observed in galaxies across all of cosmic time, though it appears AGN feedback through powerful outflows was likely most effective during the "cosmic noon", when powerful quasars were active, and SF and AGN activity peaked. Most studies focus upon this epoch to study the influence AGN feedback has upon the evolution of the host galaxy, with most analytical models predicting this feedback mode is responsible at least for the observed relationships between BH mass and host properties discussed in section 1.2.1.3.

Galaxies exhibiting feedback through this mechanism are known as "radiative mode" AGN, as the majority of their energy is emitted as radiation due to BH accretion (Heckman & Best, 2014). Theoretical models predict galactic outflows can be produced by radiative mode feedback in three different ways:

1) **Radiation Pressure-driven**: Outflows are directly accelerated by radiation pressure from the AGN on dust grains in the ISM
2) **Momentum-driven**: Outflows are driven by the coupling of the ISM and nuclear winds, but experience efficient cooling, thus conserving only their momentum
3) **Energy-driven**: Outflows are driven by the coupling of the ISM and nuclear winds and experience negligible cooling, thus retaining all of the AGN wind's mechanical energy

Radiation pressure-driven

One mechanism through which powerful, large scale galactic outflows can be generated is through the radiation pressure that AGN photons exert on dust grains embedded in the host's ISM (Fabian, 1999; Ishibashi & Fabian, 2012; Thompson et al. 2015; Ishibashi et al. 2018). Recent models predict that outflows on kpc scales should be driven by this radiation pressure on a dusty ISM (Ishibashi & Fabian, 2012, 2015; Ishibashi et al. 2018). AGN luminosity as a result of accretion is preferentially emitted as UV photons, for which dust is an efficient absorber. Thus, the dust becomes driven outwards by this radiation pressure, ionising the dust and, assuming that the gas and the dust are efficiently coupled, pulling the gas along with it. For a galaxy exposed to an AGN accreting at the Eddington limit, the effective cross-section for the interaction between the dusty gas and the radiation, in comparison to the Thomson electron scattering cross-section, is $\sigma_{\rm d}/\sigma_{\rm T}$ ~1000 (Fabian et al. 2008a, 2012). Hence, an AGN at the standard Eddington limit for ionised gas would be at an effective Eddington limit for a dusty gas ~1000 times larger. In fact, this opacity ratio remains high even for an AGN accreting below the Eddington limit: for a completely ionised medium within a few pc of the black hole, the cross section can still remain high at $\sigma_{\rm d}/\sigma_{\rm T}$ ~500, suggesting the mechanism may even be able to drive outflows for AGN accreting below $L_{\rm Edd}$ (Fabian, 2006; Fabian, 2012). Furthermore, the UV photons will be re-emitted by the dust in the IR band, so should the gas be optically thick to IR radiation, the gas, and thus also the dusty medium, could experience a momentum rate boost due to photon trapping (Roth et al. 2012; Fabian, 2012; Thompson et al. 2015).

Momentum-driven

In this scenario, a highly-ionised, sub relativistic nuclear wind is accelerated by radiation pressure due to BH accretion. These winds can be detected in observations, where they manifest in the form of either ultra fast outflows (UFOs) in X-rays (Tombesi et al. 2010), or as broad absorption lines (BALs) in UV and/or X-ray observations (Elvis, 2000). These destructive winds slam directly into the ISM, shocking and accelerating it. At this point, the gas becomes "momentum-driven" if it experiences a rapid (on timescales much shorter than its flow time) and efficient radiative cooling. This primarily occurs via inverse Compton cooling on the radiation field of the AGN (Faucher-Giguère & Quataert, 2012; Nayakshin, 2014). In this situation, the outflowing shocked gas has lost much of the pre-shock energy from the AGN wind, thus can only transfer ram pressure to the rest of the ISM, conserving momentum. This "momentum-driven" mode can only operate on scales up to ~kpc distances, above which the gas cannot efficiently cool (King et al. 2011; Costa et al. 2014; King & Pounds, 2015), and "momentum-driven" outflows have been found to be inefficient in a cosmological context,

certainly when compared to "energy-driven" outflows, by numerical simulations (Costa et al. 2014). These outflows can only inject amounts of energy into the bulge ISM comparable to ∼10% of the total bulge binding energy, thus are insufficient to destroy the bulge (King & Pounds, 2015). This mode does, therefore, present a stable environment for BH mass growth, and thus for the $M_{BH}-$host relationships to form.

Energy-driven

This mode arises in much the same way as the "momentum-driven" mode, but in this scenario the gas does not experience strong cooling. In this case, the shocked ISM retains all of the mechanical energy injected by the AGN wind. In this energy-driven limit, the outflow is much more destructive. In the "energy-driven" scenario, the momentum rate can be boosted to values much higher than L_{AGN}/c, up to a factor of ∼20 (Faucher-Giguère & Quataert, 2012, Costa et al. 2015). As such, AGN must operate in both this mode and the momentum-driven mode, else they would destroy the bulges of their hosts long before reaching observed BH masses (King & Pounds, 2015). Furthermore, models with AGN operating in the "energy-driven" mode alone cannot reproduce the observed $M_{BH}-\sigma_*$ relationship ($M_{BH} \propto \sigma_*^4$), instead returning $M_{BH} \propto \sigma_*^5$ (although recent results suggest an index more likely between 4 and 5, Shankar et al. 2016) and significantly lower normalisation parameters (Silk & Rees, 1998; King & Pounds, 2015). Hence, many models instead consider AGN to initially operate in the "momentum-driven" mode while the BH mass is insufficient to produce winds powerful enough to overcome the bulge gravitational potential, allowing the outflowing gas to re-accrete and grow the black-hole mass. Once the BH is sufficiently massive, the radiative force of the AGN photons can balance the inward gravitational force, preventing further accretion and mass growth, and establishing the observed $M_{BH} - \sigma_*$ relationship. This can be shown through a simple calculation (e.g. Fabian, 1999, 2012; Fabian et al. 2002; King, 2003, 2005; Costa et al. 2014; King & Pounds 2015), though it must be noted that this derivation of the problem is, as mentioned, simplistic in construction. The questions of which scenario provides the primary mechanism for driving outflows and the role AGN feedback plays in the establishment of the $M_{BH}-$host relationships remain key questions in Astronomy today (Peng, 2007; Kormendy & Ho, 2013; Anglés-Alcázar et al. 2017).

Finally, recent theoretical models and simulations of outflows in both AGN-driven (Zubovas & King, 2014; Costa et al. 2015; Richings & Faucher-Giguère, 2018) and starburst-driven outflows (Zhang & Thompson, 2012) have explored the possibility that outflows in this mode may radiatively cool, leading to a cold gas component able to form molecules, helping to explain observed fast molecular outflows (Feruglio et al. 2010; Sturm et al. 2011; Cicone et

al. 2014). Finally, whilst the initial temperatures of the shocked ISM are very high ($\sim 10^7$K), simulations that incorporate metal-line cooling and free-free emission (Costa et al. 2014, 2015) show clumps of colder ($\sim 10^4$K) gas can form within the outflows (Descataldo et al. 2019), with conditions appropriate for SF (positive feedback).

1.2.3 Positive Feedback

Whilst we have only previously discussed feedback in the framework of the suppression or quenching SF within a galaxy (negative feedback), the induction of SF due to feedback (positive feedback) has also been theorised. Positive feedback is predicted by a number of theoretical models to occur due to feedback from outflows compressing the molecular gas contained within the ISM, leading to the fragmentation and eventual gravitational collapse of cloud thus facilitating SF (Rees 1989; Silk & Norman, 2009; Nayakshin & Zubovas 2012; Silk, 2013; Zubovas et al. 2013a; Nayakshin, 2014; Bieri et al. 2016). More recently, however, a second channel of positive feedback has been proposed, wherein the gas contained *within* the outflows themselves should cool and fragment, providing the ideal environment for SF to occur (Ishibashi & Fabian, 2012; Zubovas et al. 2013c; Silk, 2013; Zubovas & King, 2014; Ishibashi et al. 2013; Gaibler et al. 2012; Dugan et al. 2014; Ishibashi & Fabian, 2012, 2014, 2017; Zubovas et al. 2013b; El-Badry et al. 2016; Wang & Loeb, 2018; Mukherjee et al. 2018). This section will serve to introduce the theoretical predictions and observational evidence both channels of positive feedback.

1.2.3.1 Induced SF within the galactic disk

The introduction of Integral Field Spectroscopy (IFS) has provided a way for astronomers to spatially resolve both galaxies and outflows previously only traceable through single slit spectroscopy. This has enabled the study of the effects galaxies have upon their host, and directly led to the first detections of positive feedback due to AGN-driven outflows. Several clear examples of SF triggered in shocked and compressed gas clouds have been unambiguously seen in our Galaxy (Zavagno et al. 2010a,b; Brand et al. 2011; Dewangan et al. 2012; Thompson et al. 2012; Dewangan & Ojha 2013; Deharveng et al. 2015; Ladeyschikov et al. 2015; Duronea et al. 2017; Dwarkadas et al. 2017; Figueira et al. 2017; Baug et al. 2018; Lim et al. 2018). On larger scales, evidence for SF triggered by the interaction of jets or outflows with the interstellar medium or CGM of galaxies has also been detected, although the inferred SFR is generally low ($< 1 M_\odot$ yr^{-1}; Croft et al. 2006; Elbaz et al. 2009; Crockett et al. 2012; Cresci et al. 2015b; Salomé, Salomé & Combes, 2015; Santoro et al. 2016; Lacy et al. 2017; Molnàr et al. 2017). At high redshift, indications of SF triggered by outflows or radio jets

has also been found in some cases potentially reaching a few hundred $M_\odot$ yr^{-1} (Bicknell et al. 2000; Cresci et al. 2015a).

Additionally, AGN feedback may trigger SF in objects external to their host. Specifically, SF clumps around Centaurus A (Crockett et al. 2012; Santoro et al. 2015, 2016) and induced SF in companion galaxies to AGN hosts (Croft et al. 2006; Feain et al. 2007; Elbaz et al. 2009) have both been observed, likely triggered by cold gas compression due to AGN jets.

1.2.3.2 SF within galactic outflows

Recent observations and subsequent studies of outflows have revealed that a majority of fast galactic outflows contain large amounts of cold molecular gas that likely dominates the total mass content of the outflow (Feruglio et al. , 2010; Sturm et al. 2011; Cicone et al. 2014; Garcia-Burillo et al. 2015; Combes et al. 2014; Sakamoto et al. 2014; Fluetsch et al. 2019). Much of this gas is also found to be incredibly dense ($\sim 10^6 - 10^8$ cm^{-3}) (Aalto et al. 2012, 2015; Walter et al. 2017; Privon et al. 2017) and highly clumpy (Pereira-Santaella et al. 2016; Finn et al. 2014; Borguet et al. 2012). These conditions are incredibly similar to those of HII regions in SF galaxies, and support recent models that suggest they may play host to ongoing star formation.

These models have theorised that the compression, cooling and fragmentation of the cold gas within the outflows could induce SF *within* the outflow (Descataldo et al. 2019) at rates of up to a few/several 100 $M_\odot$ yr^{-1} (Ishibashi & Fabian, 2012; Silk, 2013). Furthermore, depending on the radial location and velocity at which the SF occurs, these stars may either become gravitationally bound to the galaxy or may escape their host galaxy or dark matter halo completely (Zubovas et al. 2013c). Those hypervelocity stars that escape their progenitor could explain observations of unbound stars and SNe that populate the CGM and IGM in galaxy clusters (Ferguson et al. 1998; Gal-Yam et al. 2003), and would contribute to the observed chemical enrichment of the CGM and IGM via SNe explosions (Zubovas et al. 2013a), reducing the need to solely invoke outflows as a way to deposit the required metals into the IGM. The models predict this channel of SF should occur in outflows across all of cosmic time, and unbound young stars formed in outflows in the early times could have contributed to the reionisation of the Universe. Whilst unbound stars are predicted to have extragalactic effects, those stars that remain bound are predicted to affect their host. Due to their radial, non-planar orbits, they should contribute to the population, shape and velocity dispersion of the spheroidal component of their host (Dugan et al. 2014; Ishibashi & Fabian, 2014, 2017). Some models have invoked this contribution to account for the mass-radius relationship observed in early-type galaxies (Ishibashi & Fabian 2014) and the size growth problem of massive galaxies between today and z$\sim$2.

The study of positive feedback in local galaxies is one of the major aims of this thesis, and both the first detection of positive feedback in a single targeted source and the subsequent study of positive feedback in a blind study of outflows detected in the SDSS-IV MaNGA are discussed in depth in chapters 4 and 5 respectively.

1.3 Outflows

Many of the previously discussed feedback models require outflows as a mechanism to either quench SF by removal of gas reservoirs in the ISM, or to enhance SF in either the host galaxy or in the outflow itself. Radiative AGN feedback models (discussed in section 1.2.2.2.2) predict the highest velocity and most powerful outflows, and are required by cosmological simulations to rectify the discrepancies between the observed and predicted galaxy mass functions at the high mass end, and to reproduce the observed BH-host galaxy property relationships. Due to their nuclear driving source, AGN-driven outflows are typically found at the centres of galaxies[d], though their orientation relative to the galactic disk has been observed to often be random (Kinney et al. 2000; Veilleux et al. 2005). Conversely, stellar feedback models predict relatively lower velocity outflows, typically much less than ~1000 km s^{-1}, driven by the momentum and mechanical energy provided by SNe and stellar radiation pressure throughout starburst episodes; specifically, powerful winds from massive O- and B-type stars, and slightly later from Wolf-Rayet stars, drive outflows early into a starburst episode through radiation pressure onto the dusty ISM, before SNe explosions due to the core collapse of these massive stars dominates until the explosions of the least massive star (~10 $M_{\odot}$) at ~40 Myr. These outflows are typically invoked in galaxy evolution models to resolve the low mass end discrepancies of the galaxy mass function and, contrary to AGN-driven outflows, are typically orientated along the major axis of their host.

Classically, the majority of our observational evidence of AGN-driven outflows was traced in radio loud galaxies. In fact, evidence of outflowing gaseous material has observed across all redshifts for these galaxies, spanning from the local Universe (z<0.1, Morganti et al. 2005), out to redshifts as high as z ~4 (Fu & Stockton, 2009; Nesvadba et al. 2008, 2011). However, the recent introduction of spatially resolved spectroscopy (i.e. IFS) across the optical, IR and mm regimes has enabled detailed studies of outflows in radio quiet galaxies, which had previously remained elusive. These studies have further provided clear and concise evidence

[d]Whilst the assumption that, for a galaxy hosting an AGN-driven outflow, the outflow will be located at the centre of said host galaxy holds true for isolated systems, for merging systems it can be difficult to define a central point (e.g. for merging systems of comparable mass). During a merger, the interactions between the systems may increase accretion rates onto the (potentially passive) SMBHs of both galaxies, causing AGN activity and hence potential for AGN-driven outflows in both systems. As such, whilst these outflows will always originate from an AGN, the definition of a "central" location may be complicated.

Outflow Gas Phase	Primary Tracers	$<T_{gas}>$[K]	$<n_{gas}>$[(e^-, HI H_2) cm^{-3}]
Highly Ionised	X-ray absorption lines	10^6-10^7	10^6-10^8
Ionised	[OIII], Hα	10^3-10^4	10^2-10^4
Neutral Atmoic	HI 21cm, NaI D, [CII]	10^2-10^3	$1-10^2$
Molecular	CO, OH, [CII], H2 IR Lines	$10-10^2$	$\geq 10^3$

Table 1.1: The main observational tracers, typical average gas temperatures and typical average number density of the different phases of gas in outflows. Table from Cicone et al. (2018).

of the presence of AGN-driven galactic outflows at both low redshifts (Feruglio et al. 2010; Rupke & Veilleux, 2011; Sturm et al. 2011; Cicone et al. 2014) and high redshifts (Maiolino et al. 2012; Harrison et al. 2012, 2016; Cano-Díaz et al. 2012; Cresci et al. 2015b; Carniani et al. 2015; Brusa et al. 2015b, 2016), fostering the proposal that outflows may be ubiquitous in extremely luminous AGN (bolometric luminosities $L_{Bol}>10^{46}$ erg s^{-1}) (Woo et al. 2016).

Stellar feedback-driven outflows in SF galaxies have typically been traced by optical emission lines and, similarly to AGN drive feedback, have been observed in both local (Heckman et al. 1990, 2015; Lehnert & Heckman, 1996; Veilleux et al. 2005; Arribas et al. 2014; Cazzoli et al, 2014; Cicone et al. 2014; Heckman & Thompson, 2017) and high-redshift (Swinbank et al. 2009; Genzel et al. 2011; Newman et al. 2012; Bradshaw et al. 2013; Forster-Schreiber et al. 2019) galaxies. These outflows are typically more difficult to trace, as their signatures are often weak, and as such early studies of these outflows focused on galaxies exhibiting extreme properties, such as (ultra) luminous infrared galaxies ((U)LIRGS) experiencing massive and powerful starbursts. More recent studies, however, have utilised large, statistical samples that better represent the local galaxy population in their search for outflows (e.g. Cicone et al. 2014; Fluetsch et al. 2019).

1.3.1 The multi-scale and multi-phase nature of outflows

Despite this wealth of observational evidence for both AGN- and star formation-driven outflows, we still do not fully understand their influence over the evolutionary path of their host galaxy's properties, and cannot always rectify observations with models. This is due to both the multi-scale and multi-phase nature of both types of outflow, as well as observations typically gleaning information only on certain outflow phases (Harrison et al. 2018). For example, outflows can be detected in both molecular dusty and neutral atomic gas, as well as in both warm (10^4 K) and hot (10^8 K) ionised gas (e.g. Heckman & Thompson, 2017). Further, these phases interact differently with their hosts over a range of spatial scales, with different information on the formation and propagation of outflows being derived from sub-pc to kpc scales. Hence, to fully understand both the impact of outflows on their host galaxies, as well as their total energetics,

kinematic and geometric structures, and radial dependencies of their physical properties, we must obtain and combine information on all phases of an outflow, spatially resolved at all scales (Cicone et al. 2018).

1.3.1.1 Neutral Atomic outflows

Neutral atomic outflows may be our best evidence of SF-driven outflows. In the local universe, neutral atomic outflows are typically seen in by (U)LIRGs, characterised by extreme starburst and AGN activity, yet only occur in ∼1% of AGN (Rupke et al. 2005a; Villar Martín et al. 2014; Cazzoli et al. 2016; Sarzi et al. 2016; Perna et al. 2017; Concas et al. 2019). This has been proposed as evidence that SF powers these outflows rather than AGN (Concas et al. 2017; Bae & Woo 2018). However, it has also been suggested that neutral atomic outflows may only exist in obscured AGN that host massive reservoirs of cold gas that may be ejected by the outflow (Perna et al. 2017). When present, these outflows dominate the total outflow mass compared to that in the ionised phase (e.g. Rupke & Veilleux, 2013; Fluetsch et al. 2019).

The main evidence for neutral atomic outflows is this ejected cold gas, that presents itself as blueshifted wings of numerous emission and absorption lines across the entire electromagnetic spectrum. In the UV regime, they are traced by HI Lyα and Lyβ absorption lines, as well as numerous low-ionisation absorption features (Shapley et al. 2003; Talia et al. 2012, 2017). In the optical band, the NaI D absorption line doublet traces atomic neutral outflows (Rupke et al. 2005b; Rupke & Veilleux, 2011, 2015; Perna et al. 2017), whilst [CII] emission lines trace them in the far-IR (Janssen et al. 2016). Finally, they are traced in the radio band by HI 21-cm absorption (Morganti et al. 2005, 2016; Schulz et al. 2018).

1.3.1.2 Molecular outflows

Dependent on the driving source, molecular outflows have been observed to comprise varying amounts of the outflow's mass content. For star forming galaxies, the ionised phase has been observed to be as massive as the molecular phase (Fluetsch et al, 2019). Conversely, for AGN, the molecular phase is seen to dominate the mass of the outflow by up to 2 orders of magnitude when compared to the ionised phase (Carniani et al. 2015; Fiore et al. 2017; Fluetsch et al. 2019), or at least certainly for AGN with bolometric luminosities $L_{\rm AGN}$<$10^{46.5}$ erg s^{-1}. Above this luminosity, Fiore et al. (2017) found that the molecular and ionised phases contain similar masses, though one caveat of this study was that data of ionised and molecular outflows were obtained for different, disjoint samples, each of them with different biases. The molecular phase can have mass outflow rates of several 100-1000 M$_\odot$ yr^{-1}, though only at modest velocities of up to 1000 km s^{-1}, around a factor of 2 lower than the ionised phase. However, these observed

lower velocities for the molecular phase outflows are consistent not only with the higher masses of the molecular gas in both the energy and momentum scenarios, but also with the different proposed acceleration and deceleration processes for each phase (Carniani et al. 2015).

The bulk of molecular gas is in the form of H_2. However, whilst H_2 emission lines can be observed in the NIR and MIR bands (Dasyra & Combes, 2011; Davies et al. 2014), they only trace the hot and warm phases of the molecular gas, which is generally only a small fraction relative to the cold phase. Most studies opt to trace molecular outflows using either CO emission lines in the sub-mm/mm band (Dasyra & Combes, 2012; Cicone et al. 2014; Morganti et al. 2015; Dasyra et al. 2016) or OH lines tracing outflows in the far-IR spectra by characteristic P-Cygni profiles (Fischer et al. 2010; Sturm et al. 2011; Veilleux et al. 2013; Spoon et al. 2013) as this proxy. Molecular gas can also be traced by complex emission lines, such as HCN, HCO+, HNC (Aalto et al. 2012, 2015), which can trace the denser component of the molecular outflows.

1.3.1.3 Ionised outflows

Of particular interest in the scope of this thesis are ionised outflows. Within a typical galaxy, the systemic ionised gas component of a galaxy is traced by narrow nebular emission line profiles. In the presence of an ionised outflow, broad, asymmetric wings are present in the emission line profiles, traced by optical emission lines such as Hα and/or [NII]λ6585 for low-ionisation gas (Rupke & Veilleux, 2013; Arribas et al. 2014) or [OIII]λ5008 for medium-ionisation gas (Weedman et al. 1970; Heckman et al. 1984; Veilleux, 1991; Boroson, 2005), or by UV and optical absorption lines (Shapley et al. 2003; Talia et al. 2012; Bordoloi et al. 2014; González-Alfonso et al. 2017). These wings are often observed blueshifted with respect to the narrow nebular emission lines, with the often lack of an associated redshifted component of the outflow explained via dust obscuration in the disk of the galaxy.

The most commonly utilised tracer of ionised outflows, especially in AGN, is [OIII], as the emission cannot be emitted from high-density (10^9–10^{12} cm^{-3}) gas present in the broad line region (BLR) surrounding AGN on sub-pc scales due to it being a forbidden transition. Thus, broad wings in emission line profiles cannot be from the BLR. Ionised outflows have been observed in [OIII] emission across all epochs of the Universe, on kpc scales and with masses up to several 100 $M_\odot$ yr^{-1}. They have also provided our strongest evidence yet of negative feedback, especially in some high-redshift (z ~2) AGN galaxies, where globally high SF rates (~100 $M_\odot$ yr^{-1}), traced by systemic narrow Hα emission, are observed in the disk component, with ongoing SF in all locations except those that correlate with the path of the ionised outflow, traced by the [OIII] emission broad wings (Cano-Díaz et al. 2012; Cresci et

al, 2015b; Carniani et al. 2016). This has been suggested to be direct evidence of SF being quenched by quasar-driven outflows.

Many large sample studies of AGN have been undertaken utilising X-ray, radio and/or optical detections selection critera to locate AGN in large spectroscopic datasets, such as those offered by the SDSS. In these samples, a propensity for increased outflow occurrence (traced via broad wings in the emission line profiles of [OIII]) in correlation with AGN luminosity has been observed, and that fast (velocities up to a few $\sim$1000 km s^{-1}) ionised outflow occurrence is almost ubiquitous in AGN (Komossa et al. 2008; Zhang et al. 2011; Mullaney et al. 2013; Bae & Woo, 2014; Zakamska & Greene, 2014; Balmaverde et al. 2016; Woo et al. 2016; Harrison et al. 2016; Perna et al. 2017). Furthermore, for red, dusty, obscured sources identified as AGN through their X-ray emission, almost 3/4 hosted outflows. This suggests that ionised outflows are commonplace in dusty, obscured sources (Brusa et al. 2015a; Perna et al. 2015a; Zakamska et al. 2016; LaMassa et al. 2017; Toba et al. 2017), consistent with theoretical models of outflows driven by radiation pressure on dusty clouds in highly obscured environments (Costa et al. 2018).

1.3.2 Ionised outflow properties equations

The introduction of spatially resolved spectroscopy has enabled a number of key properties of outflows to be trivially inferred from their resolved 2D maps, such as the average velocity and radial extent of the high-velocity winds, with indirect measurements of other properties such as mass and SFR arising from 2D flux density maps. Further characteristic outflow properties, such as mass outflow rates, momentum rates and the kinetic powers of said outflows can thus be calculated, assuming a model explaining the propagation of the outflow can be established. Due to the often challenging and uncertain process of fully modelling an outflow, especially in the case of studies of large numbers of outflows, it is often best to invoke a simple model to estimate these properties (e.g. Carniani et al. 2015).

The simplest model one can choose to represent the outflow is a shell-like cloud of ejecta, contained within a cone centred on the driving source with a filling factor of 100%. In this model, one can estimate the mass outflow rate, $\dot{M}_o$, as the cloud mass, M_o, divided by the dynamical timescale of the outflow, τ_{dyn}, in turn defined as the time taken for the ionised gas in the shell to reach a distance from the outflow driving source, R_o, at an average velocity, v_o:

$$\dot{M}_o = \frac{M_o}{\tau_{dyn}} = \frac{M_o v_o}{R_o} \tag{1.3}$$

and thus also define the momentum rate $\dot{P}_o$ and kinetic energy rate $\dot{E}_o$ as

$$\dot{P}_o = \dot{M}_o v_o = \frac{M_o v_o^2}{R_o} \tag{1.4}$$

$$\dot{E}_o = \frac{1}{2}\dot{M}_o v_o^2 = \frac{1}{2}\frac{M_o v_o^3}{R_o} \tag{1.5}$$

As mentioned, the radial extent and averaged velocity can be directly estimated from the 2D velocity maps. However, due to the multi-phase nature of outflows, accurately estimating the mass from the flux maps is the largest challenge. Projection effects, resulting from the direction of the outflowing gas relative to the line of sight, affect both the observed velocity and the observed radius. However, it can be easily shown (Cicone et al. 2015) that the correction factor, averaged over the entire 4π solid angle, is 1; therefore, statistically, projection effects do not plague results, although they certainly introduce scatter.

The ionised phase of an outflow likely only contains a small fraction of the total outflow mass. The ionised gas is traced by both [OIII]λ5008 and Balmer (Hα or Hβ) emission, with no consensus drawn to which best traces the gas. As such, many previous studies used either the Hβ (Liu et al. 2013; Harrison et al. 2014) *or* [OIII]λ5008 (Cano-Díaz et al. 2012) emission lines to estimate the ionised mass, with recent studies advocating the use of both (Carniani et al. 2015). A full formalism of the determination of the ionised gas mass from the luminosity for both [OIII]λ5008 and Hβ is given in Carniani et al. 2015. In summary, the [OIII]λ5008 gas mass is given by:

$$M_{[\mathrm{OIII}]} = 0.8 \times 10^8 \mathrm{M}_\odot \left(\frac{C}{10^{[\mathrm{O/H}]-[\mathrm{O/H}]_\odot}}\right)\left(\frac{L_{[\mathrm{OIII}]}}{10^{44}\ \mathrm{erg\ s^{-1}}}\right)\left(\frac{< n_e >}{500\ \mathrm{cm^{-3}}}\right)^{-1} \tag{1.6}$$

where C is the "clumping factor", $[\mathrm{O/H}] - [\mathrm{O/H}]_\odot$ is the metallicity of the gas relative to solar metallicity, and n_e is the electron density. Determination of the metallicity of the outflowing gas is very difficult, relying on many strong line tracers to break degeneracies, and as such is often simply assumed as solar. Further, C can be reasonably assumed to be close to unity. For the Hβ emission, the gas mass is given by

$$M_{\mathrm{H}\beta} = 1.7 \times 10^9 \mathrm{M}_\odot\ C \left(\frac{L_{\mathrm{H}\beta}}{10^{44}\ \mathrm{erg\ s^{-1}}}\right)\left(\frac{< n_e >}{500\ \mathrm{cm^{-3}}}\right)^{-1} \tag{1.7}$$

The only unknown that remains that we can reasonably measure in our outflows is the electron density, n_e, which we can determine from the intensity ratio of the sulphur doublet, [SII]$\lambda\lambda$6718,6733, discussed in section 1.4.3.

1.4 Investigating the properties of gas via optical spectroscopy

Much of the work of this thesis studies the properties of both the ISM of outflow host galaxies and the ionised gas within the outflows themselves. This section serves to introduce our

investigative tools for probing said gas. Our primary analysis techniques employ a set of diagnostics available within the optical wavelength range, though some far-IR diagnostics were utilised in the initial detection of positive feedback, discussed in chapter 4. Additionally, I discuss the techniques used to determine the electron density of nebular and outflowing gas, the effects and corrections applied in the study of nebular emission lines, and the methodologies used for calculating SF rates in HII regions.

1.4.1 Excitation mechanisms in galaxies revealed by BPT diagrams

The most commonly employed technique for spectral classification of emission lines is based upon the diagnostic diagrams put forth by Baldwin, Phillips and Terlevich (1981). They suggested that some line intensity ratios could be used to determine the dominant mechanism for gas ionisation, initially proposing the bi-dimensional diagnostic diagram comparing the [OIII]λ5008/Hβ and the [NIII]λ6585/Hα line ratios. The main strengths of this scheme lie in the proximity in wavelength of the line pairs, making it reddening insensitive, and the easy accessibility of the lines in the optical spectral range in the local Universe. This diagram became initially known as the BPT diagram, though is now referred to as the [NII]-BPT diagram after the introduction of two further diagnostics put forth by Veilleux & Osterbrock (1987), comparing the [OIII]λ5008/Hβ line ratio against the [SIII]$\lambda\lambda$6718,6733/Hα and [OI]λ6306/Hβ line ratios, known as the [SII]-BPT and [OI]-BPT diagrams respectively.

One of the first demarcations of these diagrams was proposed by Kewley et al. (2001) (henceforth K01), who used population synthesis models alongside photoionisation models to argue sources that lay above their line could not be ionised solely by SF. The introduction of the SDSS provided data on large numbers of local galaxies, including the emission line ratios for the BPT diagrams, and served to make apparent the separation of the overall galaxy population into two distinct sequences in the diagrams, appearing as the characteristic "double-wing" shape. The "left wing" of the diagrams, characterised by lower values for the line ratios (due to increased Hα and Hβ recombination line emission), contains SF galaxies whose dominant ionisation source is massive O-B type stars in HII regions. In fact, the shape of this wing has been shown to be well reproduced by photoionisation models of HII regions as the bounding region of expected line ratios from gas, ionised by O-B type stars, with varying ionisation parameters and metallicities. The initial K01 demarcation line was seen to provide a clear yet too extreme cut of these populations in the [NII]-BPT diagram, with the locus below the line subsequently becoming known as the "extreme starburst" region. As such, Kauffmann et al. (2003c) (henceforth Ka03) introduced an empirically determined improvement to be used in conjunction with the K01 line, which serves to separate this region into two further loci: the

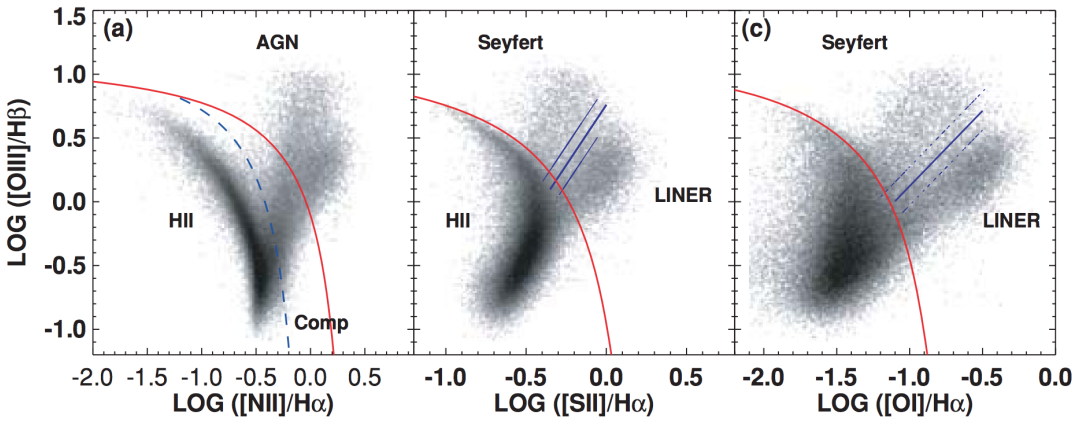


Figure 1.7: The BPT diagnostic diagrams, populated by a large sample of galaxies from the SDSS. The initial Kewley et al. (2001) demarcation line separating SF from Seyfert regions is shown in red in all panels, with the Kauffmann et al. (2003c) improved demarcation shown as the blue dashed line in the [NII]-BPT diagram (left panel). The bimodality of the right hand branch into two distinct sub-populations is most evident in the [SII]- and [OI]-BPT diagrams, with the Kewley et al. (2006) AGN/LIER demarcation line shown as blue solid lines in those diagrams. Figure from Kewley et al. (2006).

pure SF HII locus and a "composite" locus, containing galaxies experiencing ionisation from mixed sources.

Conversely, the "right wing" of the diagrams contains galaxies with various sources of ionisation, such as photoionisation by AGN, Low Ionisation Emission Line Regions (LINER) and shocks, and appear in the upper right region of the diagrams due to increased fluxes of collisionally excited lines. This wing of the BPT diagrams has been observed to exhibit a bimodality, studied in depth by Kewley et al. (2006) (henceforth K06), and generally divided into the LINER ionised lower branch and Seyfert ionised upper branch. K06 found that, from an empirical standpoint, this division between branches is most prominent in the [SII]- and [OI]-BPT diagrams. Whilst the dominant ionising mechanism in the upper (Seyfert) branch is generally considered well known to be AGN activity, the mechanisms that give rise to the LINER branch have been the subject of many studies.

Some of the first detailed studies of LINER emission argued the involved ionisation mechanisms must be significantly different from that of known active Seyfert nuclei, though the ionisation could arise from the presence of low luminosity AGN (Heckman, 1980; Ho, Filippenko & Sargent 1993). However, that same study also noted that both photoionisation from a hard radiation field (Ferland & Netzer, 1983; Kewley et al. 2006) or models of shock ionisation (Dopita & Sutherland, 1995; Dopita et al. 2015) could also produce the relatively flat power-law ionising spectrum required to produce LINER emission. In recent years, photoionisation due to

hot, evolved post-asymptotic giant branch (pAGB) stars has been suggested as a stellar source for LINER emission, due to detections of extended LINER emission on spatially resolved scales and following the same radial profile as the stellar light from the evolved stellar population (Cid Fernandes et al. 2011; Belfiore et al. 2016). However, in some active galaxies and outflows, the LI(N)ER emission can also be associated with shocks excitation (e.g. Heckman, Armus & Miley 1987; Monreal-Ibero et al. 2010)

For the purposes of this thesis, these diagnostics are applied to aid in proving ongoing SF within outflowing gas, a form of 'positive feedback', discussed in chapters 4 and 5, and to determine the likely sources driving our observed outflows and the effect of outflows on their host galaxies, discussed in detail in chapter 6.

1.4.2 WHAN diagnostics: how to tell true from fake AGN

One issue with using the BPT diagrams for classifying galaxies is the requirement for *at least* 4 lines to be detected to a significant level. For the SDSS galaxies, this left a large number of early-type galaxies unclassified. To tackle this issue, Cid Fernandes et al. (2010) introduced a new diagnostic requiring only the detection of Hα and [NII]λ6585, two lines that lie in close wavelength proximity. They proposed that the classical Hα/[NII] ratio of the [NII]-BPT diagram form one axis, with optimal transpositions of the K06 and Stasińska et al. (2006) demarcation lines serving to highlight the division between the SF loci and the more passive loci containing galaxies such as Seyferts and LINERs. For the other axis, they suggested the equivalent width of Hα (EW(Hα)) could serve as degeneracy breaking diagnostic to separate Seyfert from LINERs, from the purely heuristic standpoint that Seyfert galaxies exhibit greater EW(Hα) than LINERs. The dividing EW(Hα) was determined as 6Å, representative of the optimal transposition of the Seyfert/LINER division of K06. This diagram has become known as the WHAN diagram.

The original separations for the WHAN diagram were later expanded upon by Cid Fernandes et al. (2011) to better identify truly "retired" galaxies with no line emission from their "weak line" galaxy counterparts likely hosting weak AGN. Through a combination of empirically derived values for the SDSS sample, alongside theoretical considerations of expected Hα luminosity contributions from stellar populations of different ages and comparisons of this to the expected EW(Hα), they determined that a separating line at EW(Hα) = 3Å could separate weak AGN galaxies (above the separation, but below the classic EW(Hα) = 6Å divisor of "strong" AGN) from the "fake" AGN of retired galaxies. Whilst the diagram is not as often employed as the BPT diagrams, for many of our outflows we lack the required lines to a significant detection level for BPT analysis, and as such the WHAN provides us an economic tool to determine ionising sources in our galaxies.

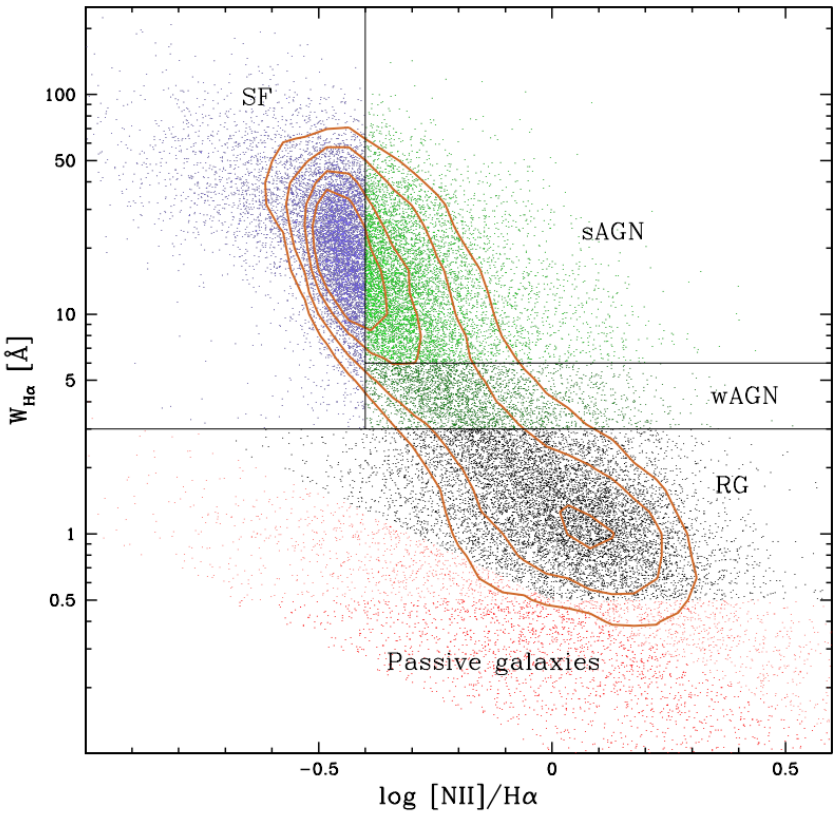


Figure 1.8: The WHAN diagnostic diagram, used to classify galaxies from their emission line properties. The demarcation lines of the diagram consist of both optimal transpositions of a number of the BPT diagram demarcation lines, as well as a separation at EW(Hα) = 0.3Å determined from stellar populations to separate "fake" from "true" AGN. Figure from Cid Fernandes et al. (2011).

Whilst an efficient diagnostic tool, the transposed Stasińska et al. (2006) demarcation line is much harsher than the line used in the traditional [NII]-BPT diagram to split HII regions into 'pure' SF and SF AGN hosts. Stasińska et al. (2006) found that in the Composite region, AGN contamination could be up to 20%, and so proposed their cut in an attempt to remove this contamination. However, objects in these regions may be fully SF, as their maximum line ratios are compatible with those expected for SF regions. Hence, the SF/AGN split in the WHAN diagram may unintentionally misclassify SF regions as AGN. Finally, similarly to the [NII]-BPT diagram, this diagnostic has a strong dependence on both nitrogen abundance (Masters et al. 2016) and the ionisation parameter (Strom et al. 2017b).

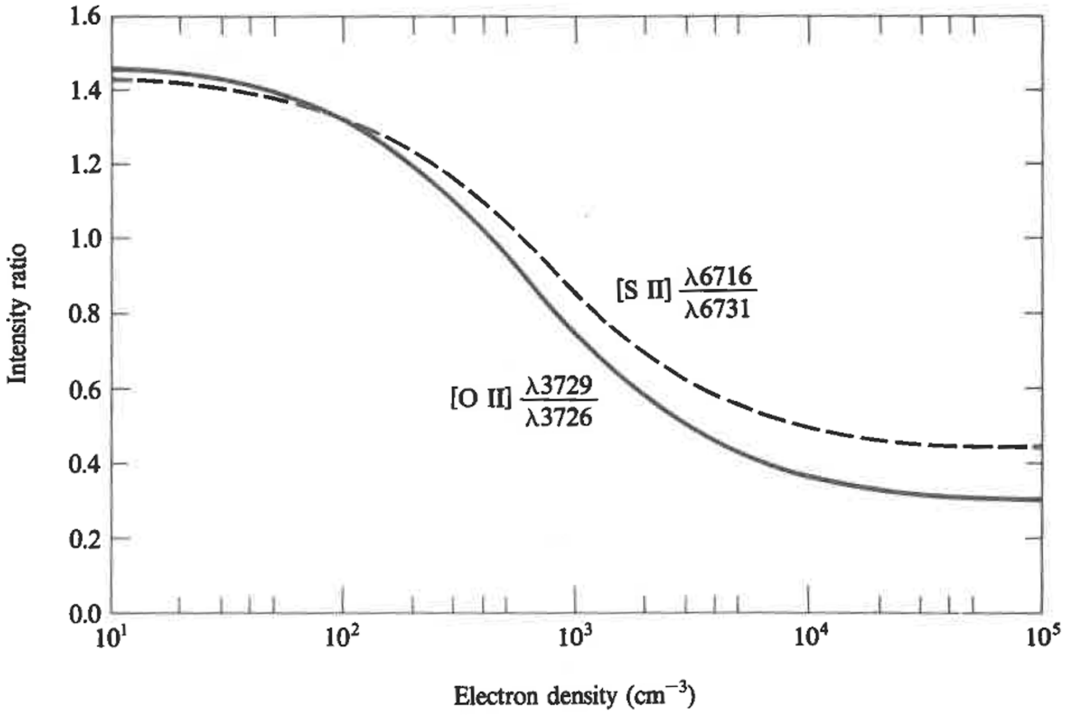


Figure 1.9: Variations of the [OII] (solid line) and [SII] (dahsed line) intensity ratios as a function of the electron density, n_e, at a typical temperature of 10^4 K. Figure from Osterbrock & Ferland (2006).

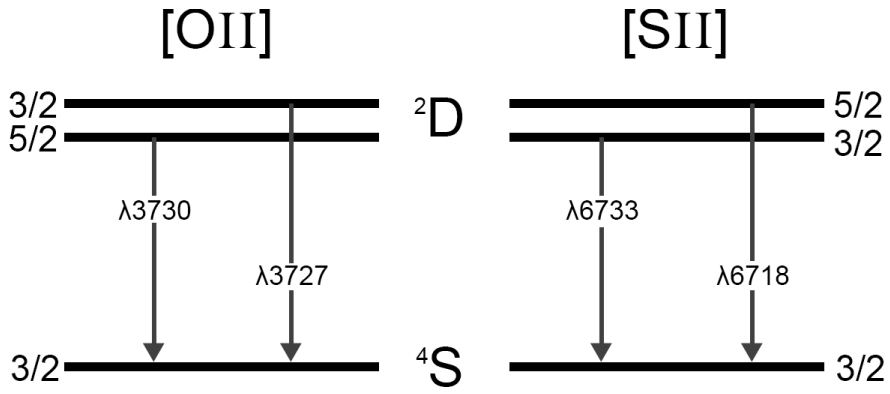


Figure 1.10: Ground state energy level structure of the O^+ ion responsible for the [OII] emission lines, and of the S^+ ion responsible for the [SII] emission lines. Shown are the two transitions for each ion that constitute the density-sensitive doublet.

1.4.3 Electron densities

One hotly debated topic within the study of outflows is on the density of the outflowing gas, with values spanning a wide range dependent on the driving source. The average electron density of the gas in a nebula (or outflow) can be measured observationally through the effects of collisional de-excitation. Typically, one considers the line emission from an ion which possesses two energy levels with similar excitation energies, where the relative excitation rates of the levels (and therefore their relative emission) depend only on the ratio of collisional strengths. If each energy level has a different radiative transition probability, collisional excitation and de-excitation rate, the relative populations of each level, and thus the ratio of intensities of the two lines, will depend on the density of the emitting gas. These ratios saturate above high- and low-density regimes, set when the density is much higher or much lower than both critical densities of both levels, and thus the line ratio is sensitive to densities within the neighbourhood of the critical densities. Within the optical wavelength range, the most commonly used lines are the Sulphur doublet [SII]$\lambda\lambda$6718,6733 and the Oxygen doublet [OII]$\lambda\lambda$3727,3730, sensitive to densities in the range 40 cm^{-3} < n_e < 10^4 cm^{-3}, typical of HII regions in galaxies. Whilst one would preferentially make use of both intensity ratios, at medium spectral resolutions, such as that of SDSS/MaNGA spectroscopy, the [OII] doublet is often blended. As such, throughout this thesis we use only the [SII] intensity ratio for determining the electron density.

1.4.4 Dust and extinction corrections

One of the key components of the ISM is dust, of which micron size grains of silicate and graphite constitute a great part (Draine, 2003). Dust affects the evolution of galaxies by facilitating the formation of complex molecules through catalysis, such as molecular hydrogen, depleting the ISM of chemical elements, and cooling ionised nebulae. Whist an in dept discussion of the role of dust is outside of the remit of this thesis, it remains important due to its attenuating affect on observed spectra. Within the optical wavelength range, dust not only attenuates light, known as "extinction" in astronomical parlance, but does so in such a way as to affect the shape of spectra, by attenuating light at bluer wavelengths more so than at redder ones, an affect known as "reddening". The extinction at any given wavelength, λ, is defined in terms of magnitudes as:

$$A_\lambda = 2.5\log\left(\frac{F_\lambda^0}{F_\lambda}\right) \tag{1.8}$$

where F_λ^0 is the unattenuated (intrinsic) flux and F_λ is the observed flux. To account for the wavelength dependence present in this definition, extinction is typically expressed in terms of

an "extinction curve":

$$A_\lambda = k_\lambda \mathrm{E(B-V)} \tag{1.9}$$

where E(B − V), known as the "colour excess", is quantified by the extinction between the B and V magnitudes.

One can determine the extinction curve of the Milky Way by observing the spectra of stars with no dust obscuration, and thus no attenuation, then locating and observing a similar star (in terms of intrinsic measurable properties) located behind significant dust layer. One of the most widely used models for extinction within the Milky Way was defined by Cardelli et al. (1989), wherein they suggest that measurements of the total-to-selective extinction in the V-band, $R_V = A_V/E(B-V)$, can parameterise the observed fluctuations around the mean extinction curve. Further studies have shown than R_V has a median values of $R_V \sim 3.1$, though it varies across lines of sight within the Milky Way (Mathis, 1990; Fitzpatrick, 1999). When one moves to observing extragalactic sources, however, this model is insufficient due to the more complex and mostly unknown distribution and geometry of the stars and dust giving rise to an "effective extinction" curve or "attenuation curve" (Natta & Panagia, 1984; Calzetti et al. 1994; Calzetti, 2001). More complex star-dust geometry will tend to provide stronger attenuation, with attenuation curves flatter than that proposed by Cardelli et al. (1989), specifically, studies such as Calzetti (2000) suggest values of R_V=4.05±0.80 are more likely for starburst obscuration. Finally, the complexities introduced by the effects of outflows on their host galaxies, and the uncertainty of their affect on dust in their host's ISM (especially in the case of radiation pressure-driven outflows, accelerated by AGN radiation pressure on the dust itself) are still not fully understood, and a field for future study.

For the work carried out within this thesis, we must correct all of our observations for the effects of reddening. The most common methodology for "de-reddening" spectra (or, more importantly for us, emission line flux measurements) is to determine the colour excess, E(B-V), from the Balmer emission lines, whose intrinsic line ratio is given by the recombination coefficients, provided one knows the extinction curve. Typically, one utilises the Hα/Hβ ratio (known as the Balmer decrement), though other ratios can be used in the absence of one of these lines. Combining equations 1.8 and 1.9, the relationship between the observed and intrinsic flux ratios is given by:

$$\frac{\mathrm{F_{H\alpha}}}{\mathrm{F_{H\beta}}} = \frac{\mathrm{F^0_{H\alpha}}}{\mathrm{F^0_{H\beta}}} 10^{-0.4\mathrm{E(B-V)}[k_{\mathrm{H}\alpha} - k_{\mathrm{H}\beta}]} \tag{1.10}$$

From theory, we known that, for Case B recombination[e], the intrinsic Balmer decrement ratio

[e]Case A and Case B are recombination situations within the theory of recombination-line radiation. For Case

is $F_{H\alpha}/F_{H\beta}$ = 2.87 at typical temperatures and election densities of T $\sim 10^4$ K and $n_e \sim 10^2$ cm^{-3} (though the intrinsic Balmer decrement depends very little on T and n_e within values typically observed in HII regions). Hence, using this intrinsic ratio as well as the Calzetti (2000) reddening curve:

$$k(\lambda) = \begin{cases} 2.659(-2.156 + 1.509/\lambda - 0.198/\lambda^2 + 0.011/\lambda^3) + R_V & 0.12\mu m \leq \lambda \leq 0.65\mu m \\ 2.659(-1.857 + 1.040/\lambda) + R_V & 0.63\mu m \leq \lambda \leq 2.20\mu m \end{cases} \tag{1.11}$$

in equation 1.10, one can determine the colour excess needed to correct observed emission line fluxes for reddening.

However, there are two considerations we must take into account when using this methodology. Firstly, in the case of emitting sources mixed with dust (more representative of galaxies) the Balmer decrement is only sensitive to modest levels of extinction. This becomes an issue when one wishes to consider the average extinction on large scales, as its value will be biased towards lower extinction values due to the domination of observed fluxes in regions of lower attenuation (i.e. typically those regions closer to us and affected by a lower column of dust). Secondly is that the stellar continuum is often found to exhibit lower reddening than the emission lines of a galaxy (Calzetti et al. 1994), thought to be due to the deeper embedding of young stars in dusty, gaseous structures.

Whilst the effect of dust must be corrected for for any calculations involving either luminosities or comparisons of line intensities at distant wavelengths, for many of the investigative tools discussed above we do not need to correct our fluxes, as the line ratios compare lines at similar wavelengths, thus affected by similar amounts of reddening (unless the lines are emitted from different regions, and thus subject to differential extinction). Finally, we also note that the extinction correction applied as such introduces one of the largest sources of error for our calculations, due to the almost 20% uncertainty on the values of R_v

1.4.5 Measuring SF rates and the Main Sequence

To quantify the level of ongoing SF in galaxies, we define the star formation rate (SFR) as the amount of mass as stars formed within a given time interval, typically in units of solar masses

A recombination, assumptions are made that all line photons emitted in the nebula can escape without absorption, thus without causing upward transitions. This situation is a sufficient approximation for gaseous nebulae that are optically thin in all HI resonance lines, and such nebulae may only contain small amounts of gas. In most cases, any galaxy with sufficient gas to be observed (such as those studied throughout this thesis) have large optical depths, hence photons will scatter with a finite probability the photons will be converted to lower-series photon. This optically thick model is known as Case B recombination and is our best approximation for most galaxies. This situation is assumed for the rest of this thesis.

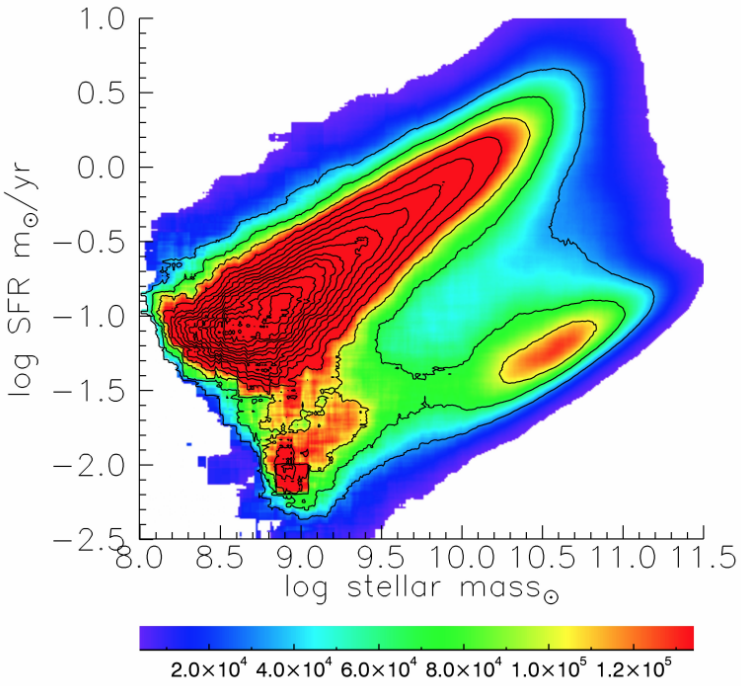


Figure 1.11: The distribution of the SDSS galaxies in the SF rate versus stellar mass plane, revealing the SF 'main sequence'. The colour-coding is representative of the density of galaxies, which has been corrected for volume incompleteness. The SF rates and stellar masses were calculated by the MPA-JHU group (Brinchmann et al. 2004). Figure from Renzini & Peng (2015).

per year ($M_\odot$ yr^{-1}). SF itself is restricted by two key factors: the amount of gas available for conversion and the efficiency with which this gas can be converted. As such, SF typically occurs in systems rich in cold, dense gas.

Many observational tracers allow us to determine levels of ongoing SF within a galaxy. The most commonly utilised tracers are near-UV (NUV) emission, which directly traces the emission of young hot stars, far-IR (FIR) emission, which traces the amount of radiation heating dust (through absorption and re-emission), and Hα emission, which serves to trace gas photoionisation by massive stars in regions of recent SF (Kennicutt, 1998; Kennicutt & Evans, 2012). Specifically, (SF) HII regions are illuminated by UV light from massive O- or B-type stars, which have typical lifetimes of only a few Myr yet emit significant amounts of ionising photos. As the Hα luminosity is proportional to the number of ionising photons in an HII region, it serves as a probe of the SFR of massive stars. As such, assuming the stellar initial mass function (IMF[f]) is well known, one can use stellar population models to define a

[f]Similarly to the previously defined mass (luminosity) function for galaxies, the IMF measures the mass distribution of recently formed stars. Salpeter (1955) measured stellar counts within the solar neighbourhood,

relationship between $L_{\mathrm{H}\alpha}$ and SFR. Kennicutt (1998) provides us with a functional form for this relationship, assuming a Salpeter IMF (Salpeter, 1955) and Case B recombination at a temperature of T$\sim 10^4$K, of:

$$SFR = L_{\mathrm{H}\alpha} \cdot 7.9 \times 10^{-42} \tag{1.12}$$

where $L_{\mathrm{H}\alpha}$ is measured in erg s^{-1} and the resultant SFR is given in units of M$_\odot$ yr^{-1}.

One dichotomy previously mentioned is the bimodality present in the SFR-M$_*$ diagram. Studies of the SFRs in local galaxies, determined using spectroscopically calculated Hα measurements from the SDSS, have shown a tight correlation exists between the SFR and the stellar mass of blue, SF galaxies (Brinchmann et al. 2004; Renzini & Peng, 2015). This correlation is parameterised by a linear relationship, that takes the form

$$SFR = \alpha M_*^{\beta} \tag{1.13}$$

where α is a normalisation factor, and β is found to be within the range $0.7<\beta<0.9$, with typical scatter around this line of the order of only 0.2-0.3 dex (Speagle et al. 2014). This relation has become commonly known as the Star Formation Main Sequence (SFMS, or more concisely MS), and applies only to SF galaxies, separated in SFR from their quiescent, red-sequence counterparts by over an order of magnitude at any given stellar mass. In fact, this separation is further exemplified when one determines the distance of all galaxies from the MS along the SFR axis, denoted as ΔMS. On the ΔMS galaxy count distribution diagram, two distinct peaks fall at ΔMS=0 and $\Delta MS \sim -1.5$, representing the MS and red-sequence galaxies respectively. Furthermore, when investigating galaxies on spatially resolved scales, this equation can only be applied to regions confirmed to be HII regions through the previously discussed BPT diagrams, as AGN or shock excitation can contaminate the Hα luminosities.

Studies that have explored the SFRs of galaxies out to the higher redshift Universe (Bell et al. 2005; Noeske et al. 2007; Daddi et al. 2007) have demonstrated both the bimodality of the diagram, as well as the existence of a MS, persist throughout the "cosmic noon" of the Universe ($z \sim 1-3$), whilst the normalisation factor α and slope of the relationship β have evolved throughout cosmic time. More specifically, going from today ($z = 0$) towards higher redshifts, the SFR at a fixed stellar mass steadily increases at a rate of a factor of about 30 (Daddi et al. 2007; Noeske et al. 2007; Rodighiero et al. 2011) peaking around $z \sim 2$ (Lilly et al. 1996; Madau & Shull, 1996). The existence of such a relationship further supports the evolutionary

restricted to stellar masses > $0.4M_\odot$, and suggested that the IMF took the form of a power law $dN/dM = \phi_{IMF}(M) \propto M^{-\alpha}$, where dN/dM denoted the number of stars per unit mass, and α=2.35. More recent studies using observations from the Milky Way suggest the IMF flattens at the low mass end (unavailable to Salpeter), and is better described by either a broken power law (Kroupa et al. 1993; Kroupa, 2001) or a log-normal distribution (Chabrier, 2003).

proposition put forth by the colour-mass diagram of a smooth, secular evolution driven by gas accretion.

Once final observation of the diagram is the existence of a group of galaxies that exist above the main sequence, commonly referred to as starburst galaxies. These galaxies are forming stars at an elevated rate with respect to normal MS galaxies, and their enhanced levels of SF are thought to be induced by galaxy interactions and merger events that serve to provide significant amounts of gas to fuel SF, and drastically reduce the depletion timescales of the interacting system when compared to the secular evolution they would experience. These starbursts could also, feasibly, be caused by compression of the gas within the galactic disk due to outflows (positive feedback).

Chapter 2

Integral Field Spectroscopy and the SDSS-IV MaNGA Survey

This chapter is partially adapted from both 'Widespread star formation inside galactic outflows', R. Gallagher et al. 2019, MNRAS, 485, 3409 and 'The properties of ionised outflows and their host galaxies', R. Gallagher et al. in prep.

2.1 Integral Field Spectroscopy

Much of our current understanding of the formation and evolution of galaxies is owed to highly successful surveys undertaken both at low redshift, such as the Sloan Digital Sy Survey (SDSS; York et al. 2000, Strauss et al. 2002), as well as at higher redshifts at a number of facilities. Although these surveys have offered us a statistical observation of how many physical processes driving the galaxy evolution vary with environment, galaxy mass and time, all suffer from the same impediment: to spectroscopically target more than a few hundred galaxies, one can only observe small subregions of each galaxy, defined by either the orientation of the slit or, as was the case for the SDSS Legacy survey, the location at which the light collecting fibre is targeted. Hence, any spectroscopically computed properties will, at best, provide one dimensional trends along a single position angle or, in the worst case scenario, provide only integrated properties for the galaxy, biasing inferred measurements due to targeting of only regions, such as galactic centres. Thus, all of the data pertaining to the rich and complex internal structure of the observed galaxies is lost.

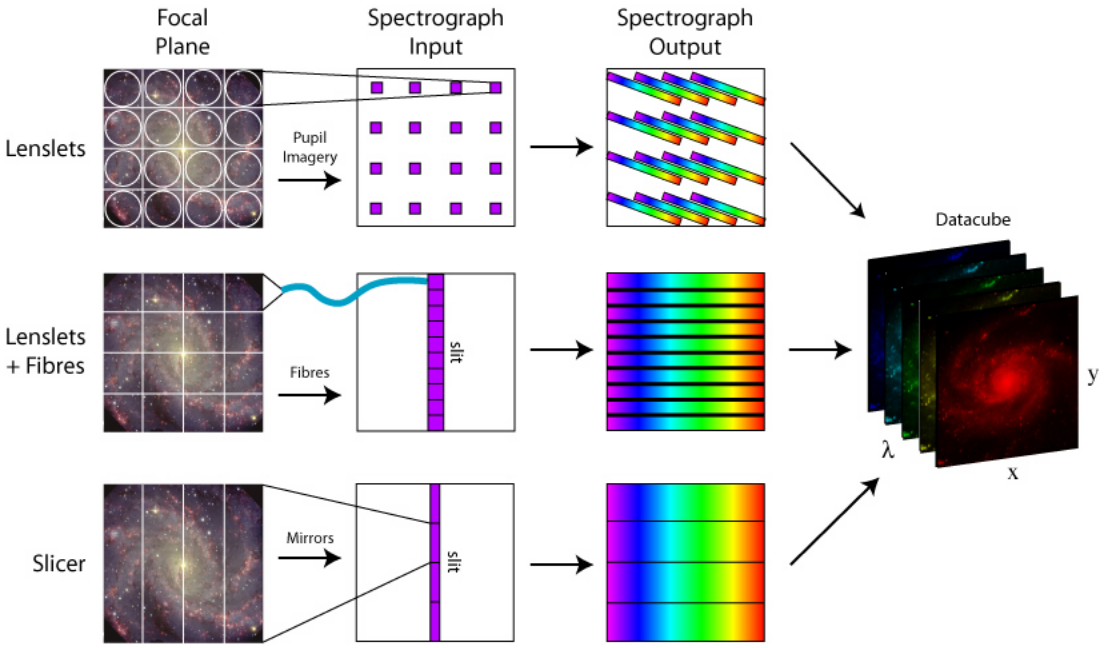


Figure 2.1: The main techniques used to achieve integral field spectroscopy. At the focal plane of the telescope, one may place either a lenslet array feeding directly into a spectrograph (top row), a combination of lenslets and fibre optic cables to feed into a spectrograph (middle row) or an image slicer using mirrors to direct areas of the image onto different spectrographs (bottom row). In the first two designs a square array of apertures are shown, but both circular (SAMI) and hexagonal (CALIFA, MaNGA) arrays of fibre apertures have been utilised. The final datacubes produced in this way provide a spatially resolved 2D image (in the x-y dimensions) at every wavelength (the λ dimension), as shown on the right of the image. Figure from http://ifs.wikidot.com/what-is-ifs, credited to M. Westmoquette, adapted from Allington-Smith et al. (1998).

Until about 20 years ago, this issue was overcome through stepping the slit positions of a long-slit spectrograph along one on-sky spatial dimension of an object. At each position, the spectra are dispersed perpendicular to the slit, providing spatial resolution along the dimension of the slit. In this way, a 3D image of the object can be built. Whilst this technique overcomes limitations, it suffers from numerous disadvantages, such as suffering from wavelength-dependent and time-dependent slit-losses due to differential atmospheric refraction (DAR) at each observation time, uncertainties in the telescope pointing (hence difficulties re-aligning the spectra) and variable seeing during the observations. To solve these issues, integral field spectrographs (spectrographs attached to integral field units (IFUs)) have been introduced, heralding the era of Integral field spectroscopy (IFS). In these instruments, the IFUs serve to divide the 2D spatial plane into a continuous array, achieved via either a Lenslet array, bundles of fibre-optic cables, or using an image slicer (all three are shown in figure 2.1). For each spatial pixel (spaxel) in the field of view, the signal is fed into a spectrograph, with the resultant

spectra stored in a datacube, containing the 2D spatial axes (x, y or RA, DEC dependent on the orientation of the observation) and the third channel containing the wavelength (or velocity) for the spectrum in each spaxel.

There have been several IFU spectrographs developed for various telescopes around the world, operating in both the optical and the near-IR. As described in the next section, the SDSS-IV MaNGA survey has enabled an enormous leap in this area. Indeed, by exploiting multiple IFUs deployed at the focal plane of the Sloan Telescope, MaNGA obtains IFU spectra of 17 galaxies simultaneously. Hence, over its years of operations, it promises to deliver the first IFS survey of thousands of galaxies, enabling a huge leap in our understanding of the properties of galaxies.

2.2 The SDSS-IV MaNGA survey: a brief overview

Within this section I aim to introduce the MaNGA survey, along with its overarching properties, with reference to the hardware, observing strategy, galaxy sample selection and initial Data Reduction Pipeline (DRP, Law et al. 2016). Whilst the MaNGA team have developed a Data Analysis Pipeline (DAP, Westfall et al. 2019), the data products of which are available to the public for all public data, I will detail in the subsequent chapter the bespoke pipeline I developed throughout my work wth the MaNGA data specifically for the detection and reduction of galaxies hosting outflows, and as such I will not focus on the MaNGA DAP in this chapter.

In comparison to other IFS surveys (SAURON (Bacon et al. 2001), CALIFA (Sánchez et al. 2012a), & SAMI (Croom et al. 2012)), MaNGA has surpassed in terms of data recovery, whilst providing data of comparable quality. When comparing MaNGA to these other surveys in terms of radial coverage, spectral resolution, and angular and spatial sampling, MaNGA displays equivalent properties, with only the red arm of the SAMI instrument providing a significantly higher spectral resolution. Where MaNGA sets itself apart, however, is in both the spectral coverage offered and the total number of objects observed.

A complete overview of the MaNGA survey can be found in a number of published papers in the literature. An initial summation of the MaNGA survey, including science goals, basic considerations of the survey requirements, instrumentation, facilities, and analysis of data observed and reduced using prototype instruments and data reduction techniques to test the full throughput of the survey is given in Bundy et al. (2015). For more precise details on the final sample design and sample selection optimisation, the reader is directed to Wake et al. (2017), whilst adopted imaging quality requirements and information regarding the final sample selection can be found in Law et al. (2015). Detailed explanations of the MaNGA instrument, expressly focusing on the design, assembly and testing of the fibre bundles are

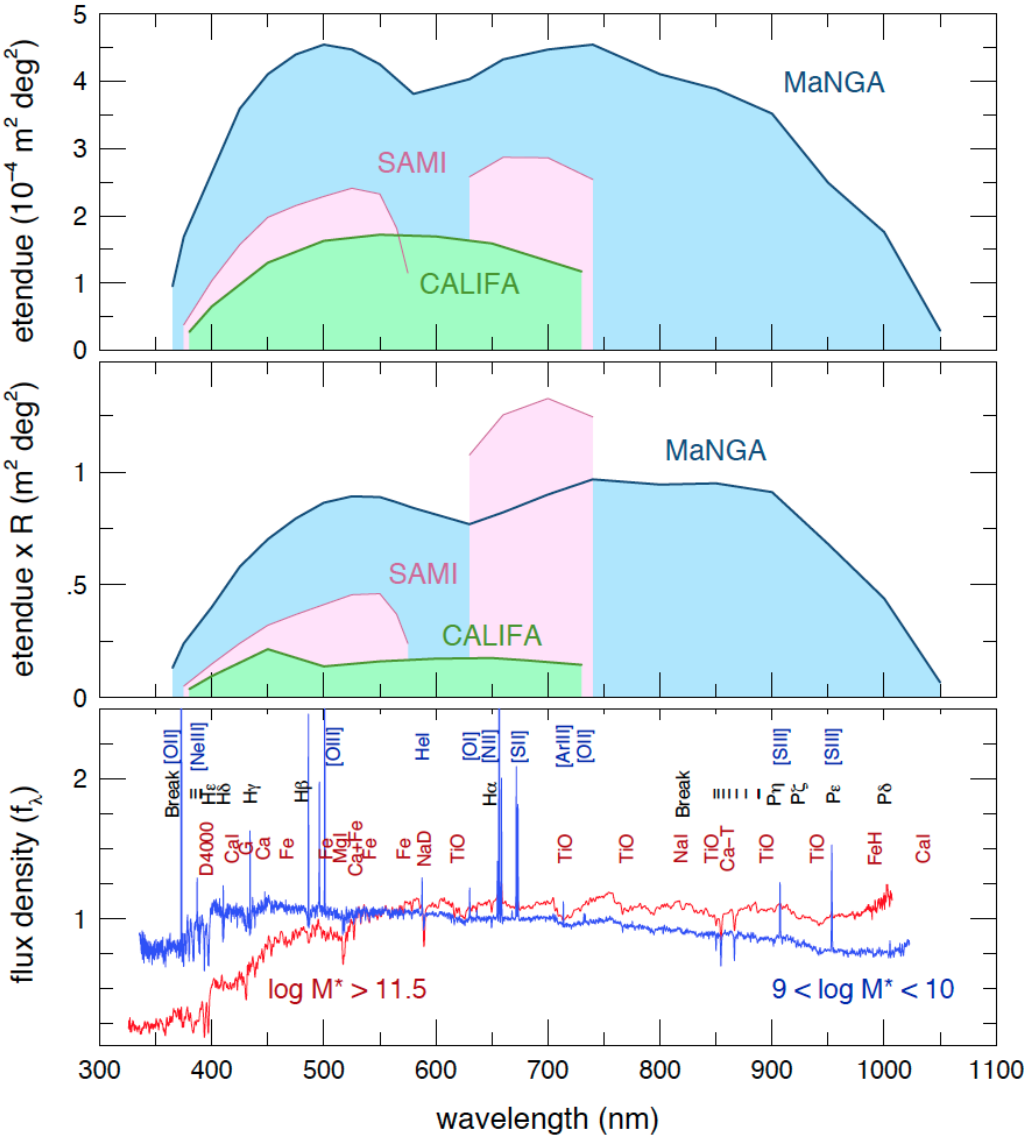


Figure 2.2: The simultaneous information gathering power of MaNGA, as reckoned by two metrics. In the top panel, the étendue (defined by the product of the system throughput, telescope collecting area and solid angle covered by all fibres) of MaNAG is shown compared to that of SAMI and CALIFA, similar IFU surveys. The middle panel shows the product of étendue and spectral resolution, showcasing how MaNGA compares favourably to these same surveys. Finally, the bottom panel shows two stacked z ~0.15 spectra from the BOSS survey, to illustrate available stellar features for stellar and ISM composition and kinematic analysis within the full MaNGA spectral range. Figure from Bundy et al. (2015).

presented in Drory et al. (2015). Law et al. (2016) present the data and software frameworks, in particular discussing development and testing of the DRP. Yan et al. (2016a) detail the methodology involved in spectrophotometric calibration, whilst Yan et al. (2016b) provide information on the execution of the survey, as well as science requirements and quality analysis of the initial data. Finally, as MaNGA has evolved over the time within which I have undertaken my projects, I have made use of two internal data releases; these are denoted by their MaNGA Product Launch (MPL) numbers. For my project 'Widespread star formation inside galactic outflows" (chapter 5) I utilised the MPL-5 dataset, equivalent to the MaNGA data publicly released in SDSS Data Release (DR) 14, details of which are presented in Abolfathi et al. (2018). Whilst undertaking and finalising my this project, there were two subsequent internal data releases: MPL-6 and MPL-7. Whilst identical in observations, MPL-7 was reduced using a newer version of the DRP, aimed at fixing issues that had been noted following the release of MPL-6. This data set is equivalent to SDSS DR 15, whose details can be found in Aguado et al. (2019). The MPL-7 dataset was used in "The properties of ionised outflows and their host galaxies" (chapter 6).

2.3 The science goals of the MaNGA survey

The past decade has seen a number of local Universe surveys showcase the power of IFS for the study of the kinematic structure, internal dynamics and radial profiles not previously accessible on a large scale. Despite the success of these surveys, they experienced a major limitation of their own: targeting libraries of only up to a few hundred galaxies, (SAURON, 72 E/S0/Sa galaxies, de Zeeuw et al. 2002; ATLAS3D 260 E/S0 galaxies, Cappellari et al. 2011). Thus, the MaNGA survey was conceived to provide the next logical step: provide the powerful IFS data required for complex studies at the statistical level of large scale surveys to truly explore the formation and evolution of galaxies in the local Universe.

The MaNGA survey was designed to answer a number of key science questions. In the context of the current ΛCDM framework, the "bottom up" hierarchical assembly of dark matter halos into increasingly massive structures provides a cosmological context for galaxy formation. Galaxy assembly and SF both drive growth throughout the early Universe, peaking around the z∼1-3 epoch. In ensuing times, a global decrease in both is seen, yet galaxy mergers and star formation, induced by inflows of gas and regulated by the feedback of various internal and external mechanisms, continue to drive growth, until the observed mass dependent cessation of Sf. The different evolutionary histories exhibited of various populations of galaxies are potentially tied to the halo and environment within which they exist, as well as the mass and gas content each galaxy itself. All of these factor into the time and environmentally dependent

morphological mixture of galaxies that have been observed. MaNGA aims to determine the physical origins of the various mechanisms at play, and their roles in driving this evolution. The key science questions put forth that the MaNGA survey wishes to address are:

1) How are galaxy disks growing in the local Universe, and what sources are supplying the gas driving this growth?
2) Within the context of the present-day growth of galactic bulges and ellipticals, what are the relative roles of major mergers, stellar accretion and secular evolutionary processes?
3) How do both internal galactic processes and externally driven processes, which in turn may depend on environment, regulate both the induction and quenching of SF?
4) Amongst different components of a galaxy, how are both the mass and angular momentum distributed, and how have they been affected by assembly through time?

2.4 The MaNGA instrumentation

Mounted on the 2.5m Sloan Telescope at the Apache Point Observatory (APO), the MaNGA IFU fibre feed system was designed with compatibility with the legacy SDSS hardware in mind. Specifically, the system utilises the spectrograph corrector optics, spectrograph-telescope mounting system for use with the BOSS spectrographs, and the full plugplate-catridge infrastructure. Succinctly, the cartridge system utilises plugplates for the positioning of optical fibres or, in the case of MaNGA fibre bundles, pluggable steel ferrules. Each plate is a circular aluminium disk, custom drilled to a targeted field at each position of an object in the 3° field of view. These plugged plates are then loaded into the cartridge, a system containing both the relevant infrastructure to attach the full system to the telescope, as well as the spectrograph pseudo-slits. The cartridge is affixed to the Cassegrain focus of the telescope via clamping to three registration ports to ensure accuracy. From this focal plane, the fibres are fed into the BOSS spectrographs. These systems have been proven cost-effective and reliable throughout the SDSS I-III projects, and are flexible enough to map the focal plane of the telescope to all of the SDSS-IV survey spectrographs, namely APOGEE-2 and BOSS.

The fundamental hardware of the MaNGA feed system is the fibre harness. Each fibre harness consists of a hexagonal bundling of fibre optic cables, and the associated mounting hardware. Although other surveys, such as SAMI, have adopted circular packing schemes due to their higher filling factor, the lower filling factor of the MaNGA fibre bundles of 56% is an acceptable tradeoff for the advantages provided by the hexagonal packing scheme, such as decreased breakages, ease of fibre replacement and simplifications to the required dithering pattern. Each MaNGA cartridge consists of 21 such fibre harnesses, organising 1423 fibres into six different configurations. The first five harnesses consist of hexagonal IFUs consisting of 19,

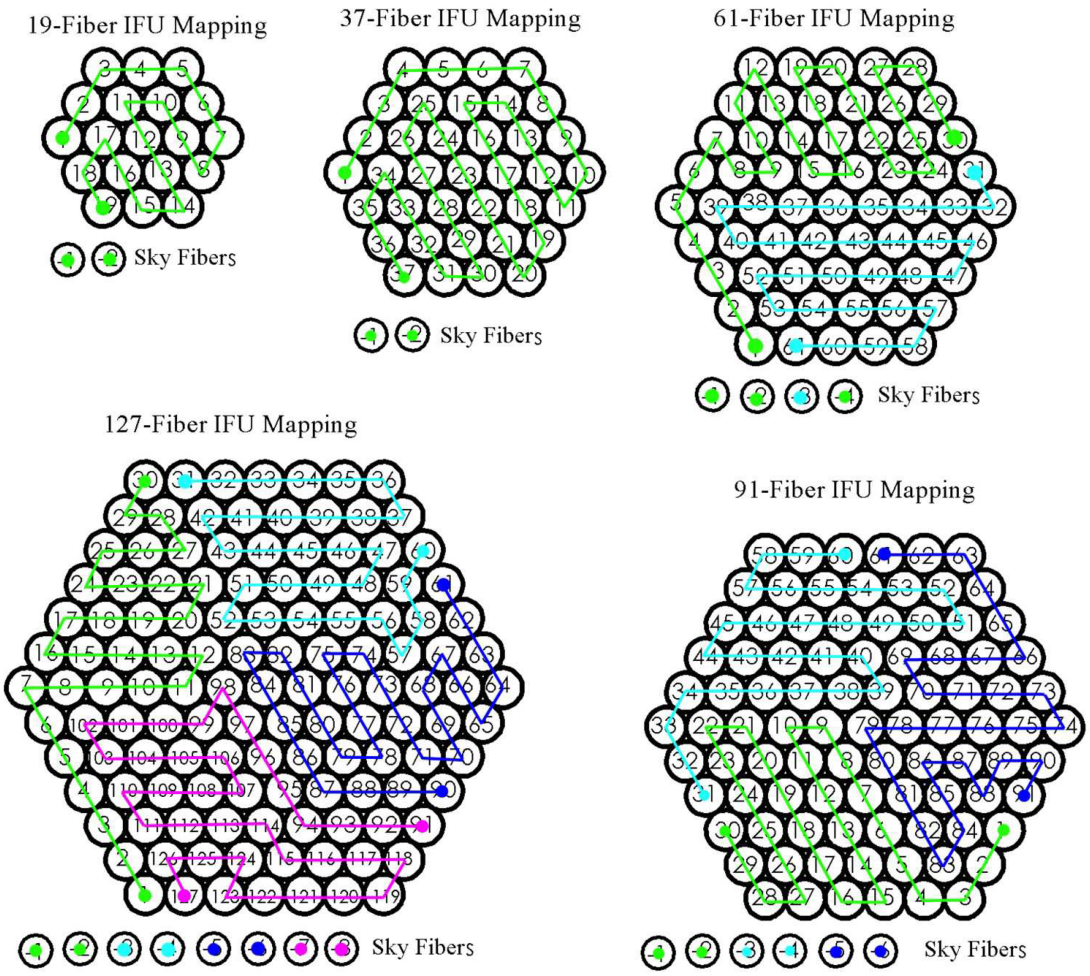


Figure 2.3: Figures illustrating both the MaNGA science IFU fibre configurations (the 7-fibre IFU used for flux calibration is not shown) as well as the mappings of the fibres from the IFU in the focal plane of the telescope to the spectrograph slits. Each colour traces fibres mapped to different v-groove blocks. Fibres with dots in them are the fibres adjacent to the sky fibres (also shown) at either end of each v-groove block. Figure from Drory et al. 2015

37, 61, 91 or 127 fibres, alongside their associated sky fibres (2, 2, 4, 6, and 8 respectively), intended for use in galaxy observations. Through optimisations of the observing strategy, with respect to the sample selection (both discussed below), it was determined each cartridge would deploy a distribution of two 19-fibre, four 37-fibre, four 61-fibre, two 91-fibre and five 127-fibre IFUs to ensure over the lifetime of the survey that fibre use was as efficient as possible. The sixth harness contains 'mini bundles' of 7 fibres, each with 3 associated sky fibres, used for spectrophotometric calibration through observations of standard stars. As such, each MaNGA observation plate consists of 121 pluggable ferrules, composed of 17 galaxy IFUs, 12 calibration "mini-bundles" and 92 sky fibres.

Each MaNGA fibre is a step-index fused silica fibre, with an internal 120μm core diameter, equivalent to 2″ at the telescope's focal plane, extending to 150μm diameter including cladding and buffers. As such, each fibre harness IFU has a diameter of between 12.5″ and 32.5″ on the Sloan telescope's focal plane, and can be deployed at any object location within the full 3° diameter of the focal plane. The fibres throughput light into the highly sensitive dual-channel BOSS spectrographs, within which the collimated beams are split via a dichroic to feed into a blue camera, covering 3600Å to 6000Å, and a red camera, covering 6000Å to 10300Å. Each camera has a wavelength varying spectral resolution, with a median resolving power of R∼2000 (R∼1400 at 4000Å & R∼2200 at 9000Å). Precise recovery of the spectral resolution is an ongoing matter for the MaNGA collaboration, with detailed analyses of variations presented in Law et al. (2016) Yan et al. (2016b). Of note, the resolution will also vary across the field of view of each observation, with resolutions at a set wavelength λ showing variance of up to a maximum of δR∼150.

2.5 Sample Selection

To determine the final number of galaxies to be included in the MaNGA sample, in such a way that uniform spatial coverage and resolution would not be compromised by the duration of the survey and the limited number of available IFU ferrules, Bundy et al. (2015) argued that a sample size of 10,000 galaxies would be a valuable and achievable goal. They consider that to provide a statistical framework from which galaxy properties could be investigated along a minimum of three "principal components" that define a galaxy population (such as stellar mass, SFR and environment), one would require a minimum of six bins per component. Typical variations between bins of these properties studied to date show precisions of 0.1 dex per decade in R_e, and thus to reach this precision to a minimum 5σ detection requires a minimum 50 galaxies per bin. Combining this information, they determine a sample size of 10,000 galaxies as a minimum requirement.

Additionally, the MaNGA sample is required to be both flat in both the stellar mass distribution and complete above a given stellar mass limit. Stellar mass is often recognised as one of the "primary components" that defines galaxy populations, hence this requirements ensures the stellar mass is adequately sampled. Further, this requirement ensures sufficiently large statistics for galaxies at the high-mass end of the distribution, which would otherwise be under-sampled in a solely volume-limited sample. With these considerations in mind, the mass range adopted by the MaNGA sample is $9 \leq \log(M_*/M_\odot) \leq 12$.

Further constraints on the sample selection arise from requirements of a minimum of 3 spatial bins for even the smallest galaxies to be observed. The typical spatial extent of galaxies

varies roughly from 3 to 9 kpc, whilst kinematic features and interesting spatial features, such as bars, bulges and spiral arms have scales roughly ∼1 kpc in size. As such, a limiting resolution of 1 kpc requires the median target redshift for a galaxy be z≈0.03. Uniform radial coverage, in terms of the effective radius R_e of each galaxy, is also required for the MaNGA sample. To achieve this, "Primary" and "Secondary" samples are defined by radial coverage goals of 1.5 R_e and 2.5 R_e (for a minimum of 80% of galaxies in each sample) respectively. The Primary sample accounts for ∼5000 galaxies, whilst the Secondary sample accounts for ∼3300 galaxies. Both are defined in an identical fashion, naturally biasing the Secondary sample to higher redshift selection limits ($\langle z \rangle$ = 0.03 for the Primary sample whilst $\langle z \rangle$ = 0.045 for the Secondary sample) and lower spatial resolutions. To the Primary sample, a "Colour-Enhanced" sample of ∼1700 galaxies are added, to account for under-representation of galaxies in rare regions of the $NUV - i$ versus M_i colour-magnitude diagram, such as high-mass blue galaxies, low-mass red galaxies and "green valley" galaxies, that trace rare yet pivotal phases of galactic evolution. The combination of the Primary sample and the Colour-Enhanced sample are known as the Primary+ sample. The final ∼1000 galaxies unaccounted for in these samples are dedicated to high-value ancillary targets.

The MaNGA selection criteria and the size distribution of the IFUs were optimised to ensure the best performance within the time constraints of the project, and are described in full in Wake et al. (2017). To ensure reproducibility, the selection strategy seeks to avoid systematic errors introduced by any "black boxes", such as codes used to calculate SFRs or stellar masses, instead making use of a set of low and high redshift limits defined by a function dependent only on i-band absolute magnitude, M_i. Further considerations for the size of the galaxy were also made, as more massive galaxies (thus higher M_i) would need to be observed at higher redshifts to ensure the required radial coverage within the field of view of the MaNGA IFUs, due to their fixed size distribution. Finally, to ensure no correlation between IFU size allocation and M_i (or stellar mass), the sample selection ensured MaNGA targeted a range of galaxy sizes at each M_i. The combination of all of these considerations led to the shape of the redshift intervals, as seen in figure 2.4.

The catalogue from which the MaNGA sample is drawn is a version of the NASA Sloan Atlas (NSA v1_0_1, Blanton et al. 2011), extended to cover galaxies to the maximum MaNGA redshift of z=0.15. The NSA is a catalogue of nearby galaxies, wherein the SDSS DR7 MAIN galaxy sample (Abazajian et al., 2009), upon which it is primarily based, is combined with additional sources to ensure redshift completeness for nearby galaxies. Of importance to MaNGA the catalogue contains both elliptical Petrosian and Sersic photomerty fits, calculated from SDSS r-band photometry, providing effective radii, as well as the axis ratio, b/a, and the angle (East of North) of the major axis. The five SDSS bands, as well as NUV and FUV

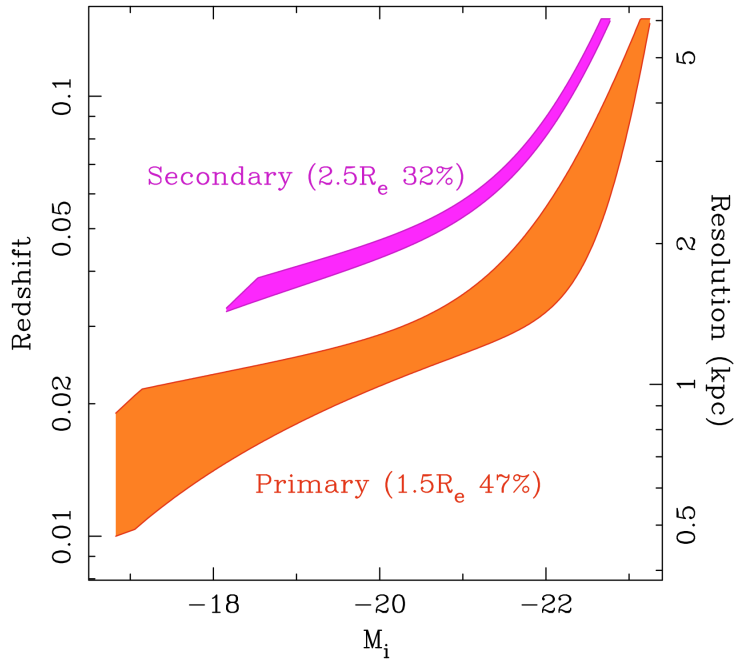


Figure 2.4: The prelimiary redshift selection cuts defining the Primary and Secondary MaNGA samples, as a function of $i-$band magnitude, M_i. Within each redshift band all galaxies are targeted, resulting in two volume-limited samples defined as smooth functions of M_i. The selection aims to provide a roughly flat distribution of galaxies in stellar mass, M_*, whilst providing uniform radial coverage out to $1.5R_e$ for the Primary sample and $2.5R_e$ for the secondary sample. Figure from Bundy et al. (2015).

bands from GALEX, with magnitudes given in the AB system and, following correction for galactic extinction, k-corrected to rest frame are also given, as well as stellar masses, computed using the kcorrect software package (version v4_2, Blanton & Roweis 2007), with Bruzual & Charlot (2003) simple stellar population models using a Chabrier (2003) initial mass function. All values calculated in the catalogue, and throughout the rest of this thesis, were calculated assuming a standard ΛCDM Cosmology, with $\Omega_M = 0.3$, $\Omega_\Lambda = 0.7$ and $H_0 = 70$ km^{-1} s^{-1} Mpc^{-1}.

2.6 Observing Strategy

One of the pressing challenges for the MaNGA observing strategy was to mitigate the under-sampling of the point-spread function (PSF) at the focal plane of the telescope due to the bundle design. The chosen hexagonal fibre-bundle configuration consists of circular apertures with large interstitial gaps due to the fibre cladding. Thus, objects that are small with respect to the fibre (aperture) size (e.g. AGN or HII regions) can appear deformed in their reconstructed

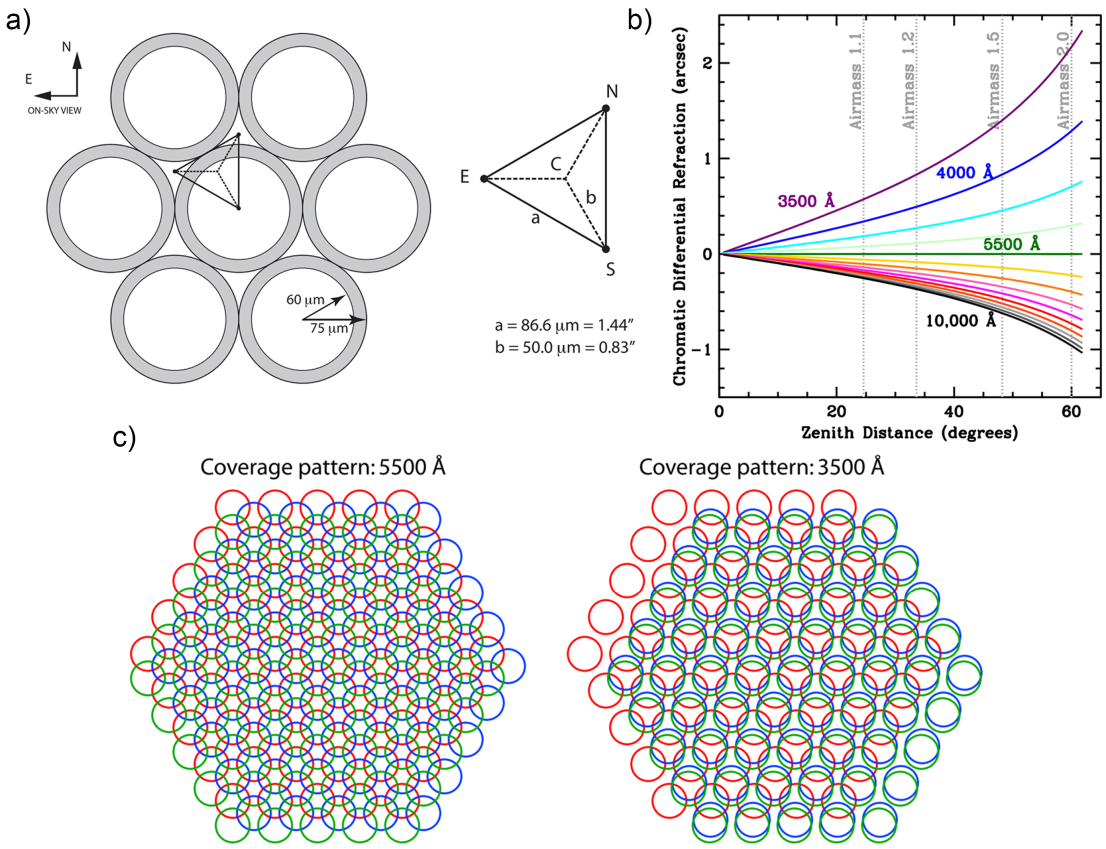


Figure 2.5: (a): Schematic design of the central 7 fibres within a hexagonally packed MaNGA IFU. The 120μm diameter of the fibre core is visible, with the surrounding cladding plus buffer. The triangle shows the relative positions of the fibre centre at each dither position. (b): The differential atmospheric refraction, in arcsec, relative to the guide wavelength of 5500Å for the MaNGA wavelength range as a function of zenith distance. (c): Degradation of the intended dither pattern due to chromatic differential refraction. Figure adapted from multiple figures from Law et al. (2015).

images, dependent of their location within the IFU and with respect to the closest fibres. To mitigate this effect, small dithers of a fraction of the fibre size are employed to sample missing points in the image plane, allowing multiple dithers to be used in image reconstruction. For MaNGA, the hexagonal IFU configuration lends itself to a triangular three-point dithering scheme, with exposures taken at each vertex, that efficiently fills the interstitial regions, as shown in figure 2.5.

Another effect that had to be taken into account in both the observing strategy and the image reconstruction performed in the data reduction pipeline is differential atmospheric refraction. Atmospheric refraction is a function that depends not only on the conditions of the atmosphere, but also zenith distance and wavelength. As such, the accuracy of the dither pattern could be degraded if each exposure within a dither set is observed at different hour angles, thus being

affected differently by atmospheric refraction. This effect is further amplified at wavelengths away from the dither guide wavelength (5500Å), due to the wavelength dependence of the atmospheric refraction. As this differential refraction can be substantial compared to the $1''$ radius of the fibres, particularly shortward of 4000Å, each recorded fibre spectrum may not sample a single region in a galaxy, and may be "bent" to other regions at different wavelengths. The mitigation of this effect is achieved for MaNGA in a two-fold manner: primarily, by the precise definition of an hour angle set, within which all exposures of a dither set must be taken (Law et al., 2015; Yan et al., 2016b); secondarily, by taking this known spectrum "bending" into account in the DRP (Law et al., 2017).

To achieve the science goals set forth by MaNGA requires each exposure to meet a minimum signal-to-noise ratio of 5 Å^{-1} fiber^{-1} int he r-band continuum at the limiting surface brightness of 23 AB arcsec^{-2}. To ensure observations are shot-noise dominated for λ>3700Å, each exposure has an integration time of 15 minutes. Thus, including the minimal overheads associated with the dithers, each dither set takes around 48 minutes. Hence, to reach the target depths, the total integration time per plate can vary between 1.5 hours to 5 hours, with an exposure time of 3 hours in median conditions (Yan et al., 2016b).

2.7 The Data Reduction Pipeline and data calibration

The MaNGA Data Reduction Pipeline (DRP) serves to take the raw frame spectra directly from APO and return sky-subtracted, spectrophotometrically calibrated spectra and rectified 3-D data cubes. The DRP runs in 2 stages: the initial "2D" stage, which produces flux-calibrated fibre spectra from individual exposures corresponding to an entire plate, and the secondary "3D" stage, wherein astrometric information is combined with the individual exposures to produce stacked data cubes for each galaxy on said plate.

The "2D" stage of the DRP performs a number of standard stages for the reduction of spectroscopic data, namely pre-processing, spectral extraction, flat-fielding, wavelength calibration, sky subtraction and flux calibration. The first pre-processing steps attempt to remove any bias from the raw exposures, from sources such as signal amplifier bias and gain factors from overscan regions. Prior to each observation, 4 second Neon-Mercury-Cadmium arclamp exposures are obtained, providing information on the fiber-to-fiber relative throughput and wavelength calibration. For each fibre, an inverse variance array is also determined from photon counts and measured read noise in each pixel, capped such that the SNR in each pixel may not exceed 100. A final stage flags cosmic rays (which significantly affect the red cameras in MaNGA) as well as issues with flat-fielding, producing a pixel mask.

Alongside wavelength calibration, the arclamp spectra are used to generate initial estimates

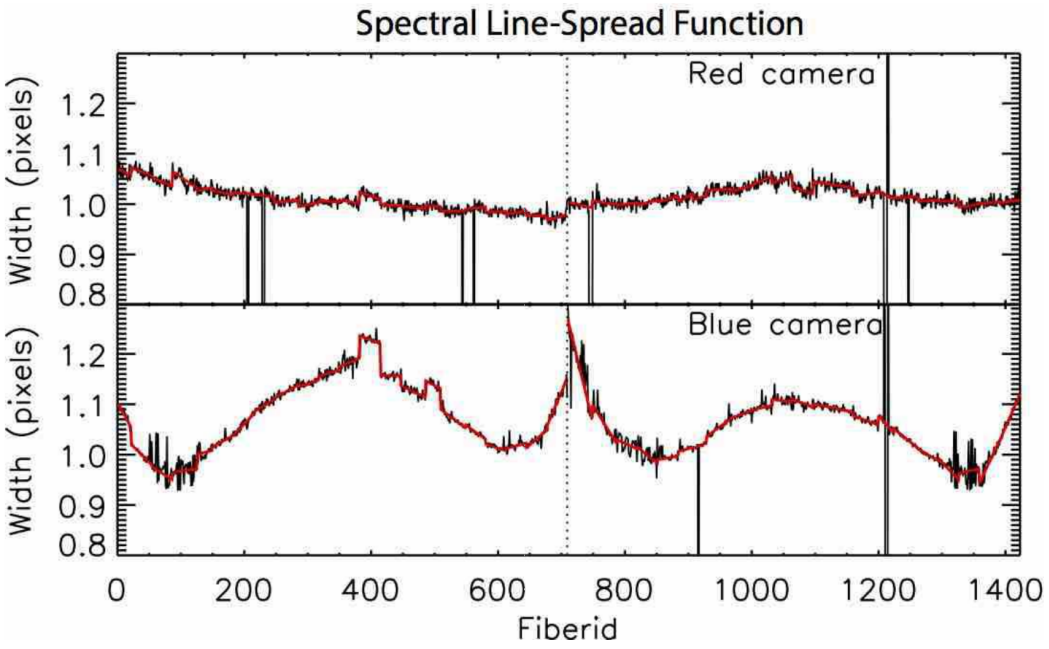


Figure 2.6: The spectral line spread fucntion (1σ LSF) of the Gaussian arclight profile, used to calibrate the spectral resolution of the blue and red cameras, as a function of fibreid. Figure from Law et al. (2016).

of the 1σ Line Spread Function (LSF) of both cameras. Using model spectra containing known emission features in the Neon-Mercury-Cadmium arclamps, the features in the observed spectra are cross-correlated and fit with Gaussian profiles, whose widths define the LSF. This LSF not only varies with wavelength, but also fibre number, and hence across the field of view of each MaNGA observation.

Next, the DRP performs sky subtraction via an iterative process making use of a super-sampled sky model constructed from all 92 sky fibres. This model is scaled to the sky background level for each IFU. This super-sampled model has a higher effective sampling of the sky than any single fibre, providing higher accuracy recovery of the LSF of sky features (e.g. OH airglow). The full procedure is detailed in Law et al. (2016). The accuracy of the calibrations and sky subtractions was tested using special "all-sky" plates, in which each IFU is placed on sky regions empty of visible sources in the SDSS catalogue. The resultant spectra are similarly calibrated to the science observations, providing an estimate to the accuracy of the procedure. The results of these observations suggest that performance is within 10% of Poissonian expectations blueward of ~8500Å, whilst errors are within 10-20% of theoretical expectations longwards of 8500Åwhere skylines are brighter.

The final steps of the "2D" stage involve flux calibration and wavelength rectification.

Whilst for previous generations of SDSS spectroscopic surveys the goal of flux calibration has been to retrieve the total flux of a point-like source (observed in a single fibre) correcting for flux losses due to atmospheric attenuation, instrumental response and flux not collected due to the fractional amount of the point-spread function that falls outside the fibre aperture, a different approach must be considered for IFU surveys. As IFU observations provide a sampling of the seeing-convolved flux profile, we do not need to make aperture corrections at the same level required for single fibre observations, thus must separate this loss from response loss. The telescope response is determined in MaNGA from the twelve 7-fibre mini-IFU bindles allocated to standard stars in each run. For these observations a model of the PSF is determined, allowing for the estimate of the relative fluxes in each fibre. This PSF model allows for aperture loss factors of the fibres to be computed and for estimation of the total flux that would be observed for each standard star if the IFU captured all of its light. Knowing the aperture loss, the system response as a function of wavelength can be determined by selecting a best-fitting template from a grid of theoretical spectra, in a similar fashion to BOSS (Dawson et al., 2013). The computed correction vectors from each star in a given exposure are then averaged to best obtain the system throughput correction needed to be applied to the science fibres. A full analysis of the flux calibration accuracy is given in Yan et al. (2016a), but in summary the calibration is found to be accurate to 1.7% between the wavelengths of Hβ and Hα and 4.7% between [OII]λ3727 and [NII]λ6585. Further, the absolute RMS calibration is better than 5% for over 89% of the MaNGA wavelength range. Finally, the flux calibrated frames from the two cameras are combined across the dichroic break onto a common fixed wavelength grid. Both decadal logarithmic and linear wavelength solutions are produced. The logarithmic solution runs between $3.5589 < \log(\lambda) < 4.0151$ with a stepsize of $\delta\log(\lambda) = 10^{-4}$ dex, whilst the linear wavelength grid runs between $3622\text{Å} < \lambda < 10353\text{Å}$ with stepsizes of 1.0Å.

Once these final flux calibrated, camera combined frames are determined for each exposure in a dither-set, the "3D" stage uses astrometry to combine the frames for each individual object into a final datacube. The astrometric calibration determines for each fibre the relative and absolute fibre location within a given IFU, offsets of frames due to location in the dithering patterns, drilled hole offsets, chromatic differential atmospheric refraction, global IFU location shift due to differential atmospheric refraction and wavelength dependent distortions. The flux and error frames are then combined with this astrometry in an image reconstruction algorithm to construct the final datacubes, with a spatial sampling of $0.5''$. Similarly to the cube reconstruction of CALIFA observations (Sánchez et al., 2012b; Husemann et al., 2013), the DRP employs a flux-conserving variant of Shepard's algorithm (Shepard, 1968), using a Gaussian weighting function.

Chapter 3

SCRAPPy - The Simultaneous n-Component Reduction and Analysis Pipeline in Python

This chapter is partially adapted from both 'Widespread star formation inside galactic outflows', R. Gallagher et al. 2019, MNRAS, 485, 3409 and 'The properties of ionised outflows and their host galaxies', R. Gallagher et al. in prep.

3.1 Introduction

Many data reduction pipelines exist today to perform often simplistic yet standard and robust data reduction techniques on raw science data from telescopes. However, for the study of atypical objects or phenomena, many of these pipelines suffer from their need to be general, and thus must be expanded upon, made more flexible and have additional functions for specific tasks added by the user. An example pertinent to this thesis is the MaNGA Data Analysis Pipeline (DAP, Westfall et al. 2019), developed to take the science data of the MaNGA DRP and provide robust data products for simple, yet accurate, analyses of all galaxies within the MaNGA sample. This required generality has lead to the inability of such pipelines to properly detect and analyse phenomena such as outflows, due to their complex emission line profiles, as well as galaxies such as those hosting type-1 AGN, where pipelines often suffer from poor fits due to their luminous central Broad Line Regions (BLRs). Due to the relative scarcity of these types of phenomena or galaxy classification (in the MaNGA sample, ~3% of galaxies

host AGN (Wylezalek et al. 2018), with a split between type-1 and type-2 AGN dependent on orientation, whlist a similar (and sometimes overlapping) fraction of galaxies host outflows (chapter 6), many of these pipelines will simply flag these objects as poorly fitted, with many studies simply opting to discard these observations from large data sets whilst able to maintain a statistically sound sample of robust and accurate derived properties.

As such, we have developed a bespoke pipeline for the express purpose of tracing outflows in IFU data: the Simultaneous n-Component Reduction and Analysis Pipeline in Python, henceforth SCRAPPy. Outside of the emission line analysis modules, many of the employed techniques are standard. Further, as this pipeline was originally developed for use with the MaNGA data (though later generalised for use with other IFU and 1D spectroscopic data, such as that of MUSE and X-shooter observations), many of the data fitting procedures and product assessment techniques were developed in parallel to the DAP development, with discussion between myself and MaNGA team members having influenced both pipelines. Certainly, for SCRAPPy in particular, these discussions were pivotal in implementing modifications to algorithms to best suit both the data, when compared to previous analyses from other surveys, and the populations to be studied. The proficiency of SCRAPPy in the detection and analysis of outflows in the MaNGA survey is showcased in chapter 5, whilst it's flexibility to work with other data sources, such as X-shooter data, was crucial to the first confirmed detection of SF within a galactic outflow, discussed in chapter 4.

3.2 Initial data analysis

SCRAPPy is designed to analyse either 1D spectra (e.g. X-shooter), 2D Row Stacked Spectra (RSS, e.g the RSS cubes of MaNGA, or stacked SDSS spectra) or 3D datacubes (e.g. MUSE or MaNGA), fed in as multi-extenstion `fits` files. Due to this desired generality, SCRAPPy utilises input files containing user modifiable data structures to both understand the input data format (e.g. the datacube extension names for required information) and direct each module's operation. In the scope of this section, the algorithms of SCRAPPy are discussed with respect to the MaNGA data.

To ensure SCRAPPy can run efficiently and encounter no issues specific to the techniques employed, an initial module has been designed to both load and analyse the input data. SCRAPPy requires that the input data file contains `DATA`, `ERROR` and `WAVELENGTH` extensions at a minimum. Whilst the `DATA` and `WAVELENGTH` extensions are assumed to be readily assessable, the `ERROR` extension may not simply be stored as the wavelength dependent 1σ array required by subsequent modules, instead perhaps being stored as inverse variances (as is the case with the MaNGA data). The onus falls upon the user to ensure SCRAPPy knows the input error format

so that SCRAPPy may convert it to a 1σ error. The wavelength-dependent spectral resolution SRES of the data may also be provided, assuming it is known. A single resolution estimate may also be given, with SCRAPPy assuming this will be a constant at all wavelengths. The SRES may be provided as either as an extension of the input cube, or as a single value or array fed directly into the SCRAPPy input wrapper.

For our analysis of the MaNGA data, our primary data source has been the DRP produced science ready datacubes, sampled logarithmically in wavelength. The MaNGA datacubes are provided in a standard format, where each cube contains 3D extensions for the DRP assessed spectra, their associated inverse variances, and a mask array containing data quality assessment flags from the DRP. Additionally, each file provides 1D arrays containing the central wavelength for each spectral channel. Dependent upon the data release, each cube will also contain either a 1D array containing the median spectral resolution for all fibres within the IFU (MPL-6 and previous releases) or a 3D extension containing the spectral resolution in each spaxel (MPL-7 onwards).

An initial assessment is performed to ensure that, for each spaxel, there is sufficient viable data to be fitted. Each spaxel has an associated mask array, containing flags in each wavelength channel defined by the DRP using the MaNGA bitmasks structure, that inform the observer of which spectral channels to exclude. Entire channels can be masked for reasons such as broken fibres, known foreground contaminants (such as stars), detector artefacts or, in the majority of cases, that the spaxel falls outside of the hexagonal MaNGA field of view. We convert each mask into a binary array that simply defines a spaxel's wavelength channel as either masked or unmasked. Whilst this ensures the majority of the data will be of sufficient quality, we must also ensure that, for each spaxel, the spectral fluxes and inverse variances are finite, non-zero values. The fitting algorithms we employ require all data (and errors) be positive, finite values. As such, whenever we encounter an invalid data (or error) value, we replace them with a sufficiently small value (e.g. 10^{-50} for the fluxes or 1 for the error arrays), adding the wavelength channel to the binary mask array so that the channel can be discarded by the fitting process. Once this inspection has been completed over the entire field of view, we collapse the resulting 3D binary mask cube into a 2D mask, containing the fractional amount of masked data to unmasked data in each spaxel. We use this to initially exclude any data for which more that 20% of the data is masked.

Next, we have to ensure the spectral resolution array for each galaxy is continuous and valid in all wavelength channels. The spectral resolution in each channel is is assumed to be representative of the width of a Gaussian line spread function (LSF). As an example the BOSS spectrograph, used in MaNGA observations, splits the collimated beams of light into separate blue (3600Å < λ < 6300Å) and red (5900Å < λ < 10300Å) channel cameras using a dichroic

beamsplitter; the resolution array represents a marriage of the resolutions of both spectrographs. Either side of the overlapping wavelength channels of the dichroic, the resolution is array is representative of only one of the spectrographs and, as such, should be smooth and continuous. Hence, any invalid channel (identified by a non finite number or a sharp gradient change between channels) is replaced by a cubic-spline interpolated value. Within the overlapping wavelength channels of dichroic however, the resolution is not smooth, and is representative of the second moment of the distribution defined by the sum of the Gaussian LSFs of each spectrograph. Within this region, we use a simple linear interpolation to replace any invalid values. The type of interpolation to use in certain spectral regions is defined by the user in an array, of length equal to the spectral resolution array.

Finally, SCRAPPy will de-redshift each galaxy before further processing. For the MaNGA galaxies, the redshift is provided by the NSA catalogue ($z_{\rm NSA}$) from which the sample was drawn, as discussed in chapter 2. This is done to prevent errors in calculating the Doppler velocities from observed redshift changes at later points in SCRAPPy. Typically, one would fit an observed spectrum or emission line at its observed wavelength, λ_{obs}, and thus attain an observed redshift, z_{obs}, according to

$$\lambda_{obs} = \lambda_{rest}(1 + z_{obs}) \tag{3.1}$$

where λ_{rest} is the rest frame wavelength of the feature that has been fit. However, this assumes no distinction between cosmological redshift due to expansion and Doppler shift due to the galaxy's internal velocity motions, with both summarily represented by z_{obs}. If one does intend to make a distinction between cosmological redshift due to the Hubble flow and Doppler shift due to rotation in a galaxy, one would first de-redshift the spectrum by the Hubble flow redshift, z_H then calculate the velocity from

$$\lambda_{obs} = \lambda_{rest}(1 + z_H)\left(1 + \frac{v}{c}\right) \tag{3.2}$$

As such, assuming we will recover the true velocity from equation 3.2, the difference between the true velocity and that obtained from equation 3.1 can be determined by substituting equation 3.2 into 3.1 and re-arranging, obtaining:

$$v = c\frac{z_{obs} - z_H}{1 + z_H} \tag{3.3}$$

This suggests a correction on the true velocity proportional to the Hubble redshift of the galaxy when compared to that recovered simply from the observed redshift. Hence, all spectra are de-redshifted, assuming the provided redshift is the Hubble redshift (z_{gal}=z_H) before any fits are performed. For the MaNGA galaxies, this correction can be as much as a 15% correction at the maximum redshift, with a typical correction of 3% at the median redshift of the Primary sample.

3.3 Spatial Binning

The main philosophy behind the design and implementation of SCRAPPy is to recover robust kinematic properties for the ionised gas not only associated with virial motions, but also those tracing outflows. We trace these two largely independent kinematic features through a set of nebular emission lines. Despite the focus of SCRAPPy, and thus this thesis, being on emission line properties, accurate measurements cannot be recovered without also accurately recovering the underlying stellar continuum, including the absorption features often present at the same wavelength as many of the nebular emission lines. The stellar component also further serves to provide a stable reference for both the velocity field and velocity dispersion of the galaxy being analysed.

Accurate removal of the stellar component has plagued large spectroscopic studies, such as the SDSS (Tremonti et al., 2004; Oh et al., 2011), since their conception, and the difficulties therein have only been compounded by the spatial dimension introduced by IFS data. Analysis of both the stellar continuum and emission lines in the faint regions in the extremities of galaxies provided by the additional spatial dimension is often difficult and prone to poor recovery of kinematics. Analysis of these issues, as well as further consequences of this additional dimension have been previously explored for prior IFS surveys, such as ATLAS3D and CALIFA, and much of the discussion contained within pipeline development papers for these surveys have significantly aided in the development of SCRAPPy.

3.3.1 Spectral Signal to Noise

For spatially resolved IFS data, it has become common practice to bin neighbouring spaxels to ensure a minimum Signal-to-Noise Ratio (SNR) for each spectrum, though this comes at the cost of loosing spatial information. With the consideration for the large variances of galaxy surface brightness profiles, typically as a function of radius, many binning schemes have been adopted, each providing different advantages and disadvantages dependent upon the science goals of the study. The most commonly adopted for IFS surveys has been the adaptive spatial-binning scheme implemeted by the Voronoi binning algorithm (Cappellari & Copin, 2003), due to its ability to optimally generate bins that properly tesselate regions of the galaxy on the sky plane, whilst remaining as "round" as possible, with uniform regularity in the SNR scatter of the bins around a defined minimum SNR. The scheme was initially designed for analysis of IFS data of early-type galaxies (Emsellem et al. 2004; Sarzi et al. 2005), for which bins defined by the continuum SNR were used for recovery of both stellar features and emission lines.

Whilst proven sufficient for these analyses, the scheme encounters issues with galaxies exhibiting irregular, clumpy surface brightness profiles, such as mergers or, more expressly of

interest to us, outflow host galaxies. Specifically, work in the literature has shown that, in early-type galaxies, the emission line surface brightness profiles follow those of the continuum (Sarzi et al. 2010; Papaderos et al. 2013; Belfiore et al. 2016), yet recent studies have shown that line emission can instead be clumpy, with disparate brightness profiles relative to the continuum. Furthermore, line emission can be very weak and diffuse, a problem often encountered at the outskirts of galaxies and in outflows. As such, any binning scheme that focuses solely upon constructing bins from the SNR of a single component of a galaxy, such as the continuum, may introduce bias by including spaxels of diffuse regions in bins constructed from bright region spaxels, with the intention of increasing the SNR in the bin but both mixing regions of vastly different properties unnecessarily and loosing crucial spatial information.

Conversely, to fully recover the emission line properties, one requires an accurate model of the stellar continuum. This issue arises from the coincidence of emission lines with both stellar features, most evident in the blue regions of the spectra that contain both weak emission lines and many deep stellar atmosphere absorption features; for instance, the Blamer emission lines lie above the Balmer stellar absorption features. As such, sufficient SNR is required in the stellar continuum for accurate extraction of, at the very least, the first two components of the stellar line of sight velocity distribution (LOSVD). This required minimum level of SNR cannot simply be quantified for all surveys: a combination of the uniqueness of each telescope throughput and the techniques employed in the algorithms to return science ready data require considerations be made for each survey undertaken. Hence, we have designed SCRAPPy to allow any non-zero, positive, finite value as the minimum SNR for the binning module. Furthermore, to remove data with insufficient signal, thus leading to large, spatially unresolved bins, from subsequent analyses, SCRAPPy can remove any data with an SNR below a threshold minimum.

Much exploration into the minimum required continuum SNR for accurate recovery of the stellar velocity dispersion, σ_*, for the MaNGA data has been performed (e.g. Westfall et al. 2019). As such, to maintain consistency with the DAP output for product comparison, we chose a minimum SNR of 10 per spectral pixel within the SDSS r-band for our analyses of the MaNGA data, sufficient for accurate recovery of σ_* for good emission line recovery, whilst ensuring bins remain sufficiently small to minimise the loss of spatial information. As an aside, the majority of our outflow spaxels happen to exist in nuclear regions of galaxies, where the (stellar) SNR is sufficiently high that no binning is needed. However, for the few extended outflows we detect the binning will affect the diffuse gas at the extremities.

In SCRAPPy, the standard procedure we employ for the calculation of the stellar SNR for each spectrum considers the flux and 1σ error within a wavelength window. However, SCRAPPy allows for custom user-defined SNR calculation algorithms that rely on any of the default input information (`DATA`, `ERROR`, `SRES` and `WAVELENGTH`). For the MaNGA data, we

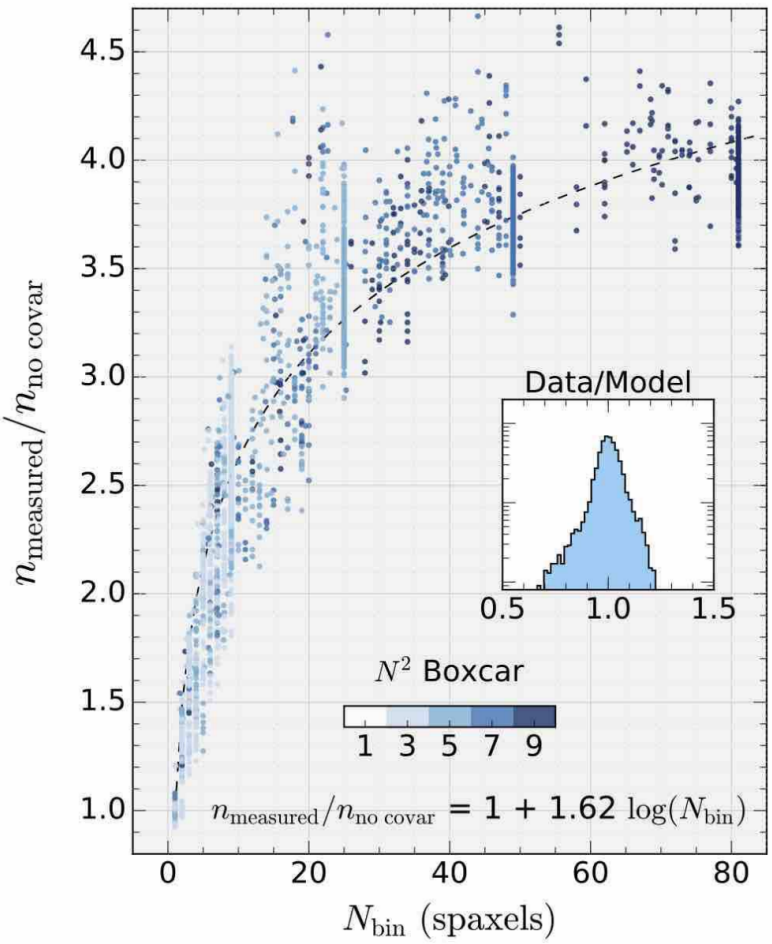


Figure 3.1: The ratio of the measured noise for a simulated MaNGA datacube, $n_{measured}$, to a calculation of noise with no consideration to the covariance, $n_{no\ covar}$, as a function of the number of spaxels in the bin used to calculate the combined spectrum. Shown as the dashed line is the empirical equation for covariance estimation for MaNGA. Figure from Law et al. (2016).

consider the values within the SDSS r-band wavelength range (5380Å < λ < 7040Å), taking the median flux value as the signal and the median 1σ error, calculated from the median inverse variance as $\varepsilon = \sqrt{IVAR^{-1}}$, as the noise for each spaxel. Using the resultant 2D SNR map, we instruct SCRAPPy to make an initial SNR cut of 3 to the continuum to ensure minimal bin sizes, as bin sizes may become too large if low SNR bins are included, diluting the signal whilst increasing noise.

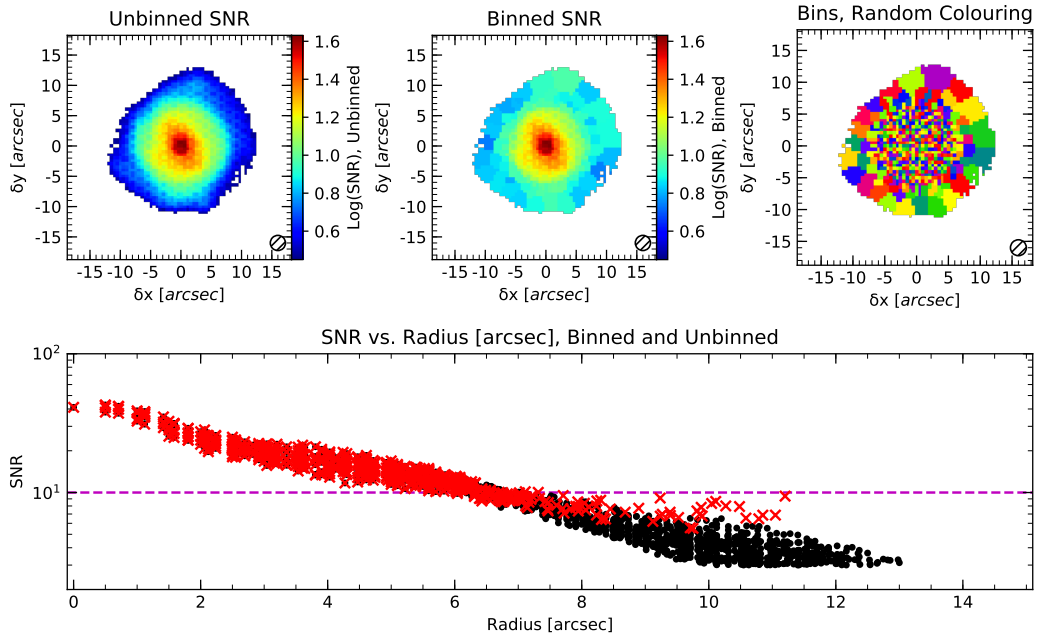


Figure 3.2: Voronoi binning of a MaNGA galaxy to a target bin minimum SNR of 10 on the stellar continuum. The top left panel shows the unbinned SNR. The middle panel shows the bins determined by the Voronoi binning algorithm, coloured by the bin SNR (with covariance included). The right panel shows a random colouring of the bins to better show the individual bins. The bottom panel shows the SNR of the individual spaxels (black points) and bins (red crosses) as a function of radius, The magenta line shows the SNR=10 target minimum SNR.

3.3.2 Voronoi Binning

The main algorithm used by SCRAPPy for data binning is the Voronoi binning algorithm `VORBIN` (Cappellari & Copin, 2003). By default, SCRAPPY utilises the standard `VORBIN` algorithm. Whilst this algorithm is sufficient for data showing little to no covariance between spaxels, such as for IFUs that employ square lenslet arrays to split the on-sky image (e.g. SAURON (Bacon et al. 2001) or MUSE (Laurent et al. 2006)), surveys that employ fibre bundle IFUs may experience covariance between spaxels due to cube reconstruction algorithms. As such, SCRAPPY utilises the `sn_func` keyword of the `VORBIN` package to allow the user to provide a different algorithm to define bin SNRs.

As an example, the data construction scheme used to produce the MaNGA datacubes (Law et al. 2016) follows the methodology of Shepard (1968) and leads to covariance between adjacent spaxels. This covariance has little affect in signal calculations, but results in noise calculations that often fail to properly account for the covariance are significantly underestimated. Calculation of the full covariance matrix is expensive in both time and computational power,

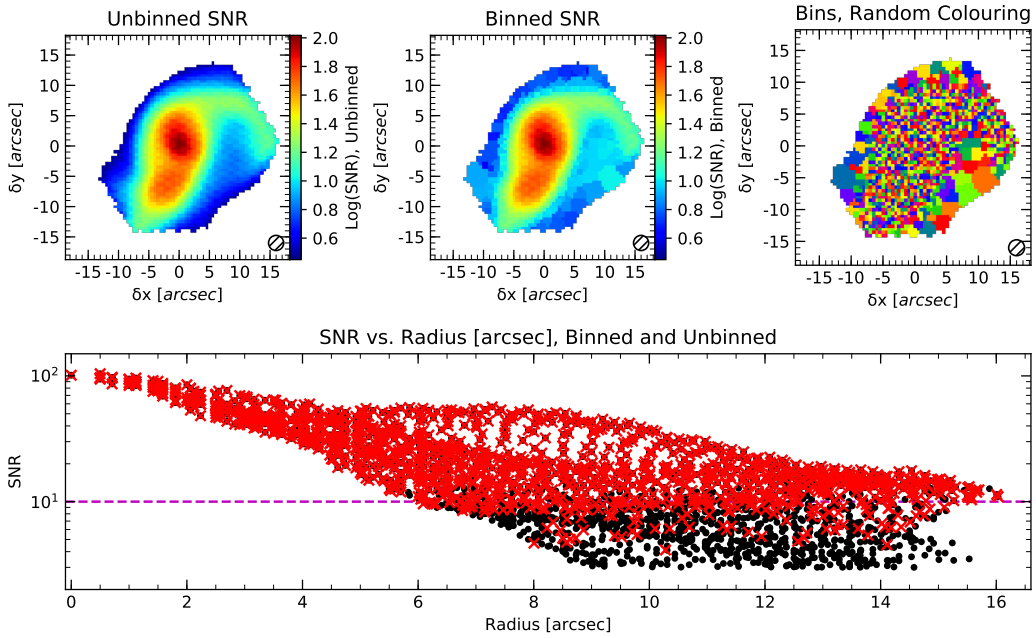


Figure 3.3: The same as figure 3.2, except for a merging system. This is shown to highlight the lack of binning in the extended arms of one of the galaxies due to the modifications made to the voronoi binning algorithm to automatically assign all spaxels with sufficient SNR to their own initial individual spaxel bins.

and has again been studied extensively for the MaNGA data. Law et al. (2016) have found that the correct noise vector can be obtained by recalibrating the propagated error vector via a simple fitting function, of similar form to that of Husemann et al. (2013), of:

$$\beta_{bin} = n_{measured}/n_{nocovar} \approx 1 + 1.62\log(N_{bin}) \tag{3.4}$$

where $n_{measured}$ is the true noise vector, $n_{nocovar}$ is the noise vector computed through propagation assuming no covariance (i.e all spaxels are independent) and N_{bin} is the number of spaxels forming the bin. The covariance thus has the effect of drastically increasing the number of bins required to reach the target SNR, an example of which is shown in Westfall et al. (2019), figure 9. Hence, we employ this empirical correction in the SCRAPPy binning module when reducing MaNGA data.

A modification to the standard procedure for using `VORBIN` algorithm can also be applied in SCRAPPy. Typically, all data would be passed to `VORBIN`, with spaxels with sufficient SNR forming single spaxel bins, and other spaxels being accreted into bins. In subsequent stages in `VORBIN`, the spaxels of bins which do not meet the target SNR are accreted into nearer bins that have met the required SNR, potentially including some of the single spaxel bins. However,

this may not always be desirable. In the case or irregular systems, such as mergers, amid lower luminosity areas in the extremities some single spaxels may have sufficient SNR to form single spaxel 'bins', yet could be conglomerated into larger bins in the traditional algorithm; thus, the rich spatial information in these key areas may be lost. To maintain as much spatial information as possible, SCRAPPy allows the user to decide whether to remove data with an SNR greater than the target SNR from the passed array prior to binning, as such only allowing `VORBIN` to bin the data with insufficient SNR. For our analyses of outflows in the MaNGA data we utilised this modification. Examples of the resultant binning can be seen in figures 3.2, which shows an outflow host galaxy with no environmental interactions, and 3.3, which shows a merging system. In the merging system, a number of single spaxel bins in the North-West corner of the plot would have been included in larger bins without the modification discussed, losing the spatial information there.

In later modules, SCRAPPy can (if desired by the user) disassemble these bins into the constituent spaxels for deeper analysis of the spatial components of the analysed galaxy. The primary use of binning in SCRAPPy is to provide accurate initial estimates to the later stage fitting modules (e.g. the individual spaxel spectra fitting) as all of the implemented algorithms utilise Levenberg-Marquardt alogirthms, implemented as the Python distribution of `mpfit`, which can be biased by initial estimates, reaching local minima rather than global minima in noisy data. Providing an estimate near to the global minimum instead will allow the algorithms to settle into these minima, rather than nearby local minima. Furthermore, this binning scheme is applied to the stellar component for better estimation of the stellar continuum yet, as discussed above, this scheme may not best recover the emission line components.

3.3.2.1 Spectral stacking within spatial bins

For each bin, our spectral stacking module must produce the flux density, 1σ error, mask and wavelength-dependent spectral resolution arrays for use in subsequent modules. A number of in built standard procedures are used by SCRAPPy that will, for most data, produce accurate, stacked data within common spatial bins. However, due to the variability between surveys and data types, SCRAPPy again allows for custom user formulae to be input for spectral stacking. Here, I present both the standard formulae used by SCRAPPy, as well as the customisations needed for the MaNGA data.

The standard procedure in SCRAPPy for calculating the flux density simply sums the contribution of all spectra contained within a bin in each wavelength channel, with respect to the mask array. Namely, for a bin i, constructed from N_i spaxels, we calculate the binned flux density spectrum as:

$$F(\lambda)_i = \frac{\sum_{j=1}^{N_i} F(\lambda)_j \phi(\lambda)_j}{\sum_{j=1}^{N_i} \phi(\lambda)_j} \tag{3.5}$$

where $\phi(\lambda)_j$ is the mask of spaxel j in wavelength channel λ, and will be either 0 or 1 for a bad or good channel respectively. The denominator ensures the flux density remains per spaxel whilst not reducing the total flux in channels with masked fluxes. The summed mask array is simply defined by the sum of all constituent masks in each wavelength channel, divided by the number of spaxels in the bin. In effect, this informs us of the fractional amount of "good" data that each binned spectrum contains. Within SCRAPPy, one can choose the fractional amount to below which channels will be masked, hence creating a binary wavelength-dependent mask of good or bad pixels. Within SCRAPPy, both the fractional and binary masks are stored. For our analysis of the MaNGA data, we used no modifications to this formula, but required each channel consist of only "good" data, i.e. $\sum_{j=1}^{N_i} \phi(\lambda)_j / N_i = 1$.

The default stacking procedure for the error (noise) arrays in SCRAPPy uses simple error propagation formulae. Similarly to the flux density calculations, SCRAPPy determines the contribution of each error spectrum contained within a bin in each wavelength channels, with respect to the mask array, summing them in quadrature:

$$\varepsilon(\lambda)_i = \sqrt{\frac{\sum_{j=1}^{N_i} \varepsilon(\lambda)_j^2 \phi(\lambda)_j}{\sum_{j=1}^{N_i} \phi(\lambda)_j}} \tag{3.6}$$

However, in the case of the MaNGA data, we must account for the covariance between neighbouring spaxels in this calculation. As such, we provide SCRAPPy with a slightly modified algorithm for the error spectum calculations:

$$\varepsilon(\lambda)_i = \beta_i \sqrt{\frac{\sum_{j=1}^{N_i} \varepsilon(\lambda)_j^2 \phi(\lambda)_j}{\sum_{j=1}^{N_i} \phi(\lambda)_j}} \tag{3.7}$$

where β_i is the empirical covariance correction value calculated for bin i, using equation 3.4.

Finally, we consider the spectral resolution arrays. We define the spectral resolution as $\mathrm{R}(\lambda) = \lambda / \Delta\lambda(\lambda)$, where $\Delta\lambda(\lambda)$ is the full-width at half maximum (FWHM) of the spectral resolution element at wavelength. In the case that a 3D wavelength dependent resolution cube is provided, SCRAPPy determines the line-spread function (LSF) of a binned spectrum from the determined LSFs of each spaxel in the bin. Namely, for spaxel j with wavelength dependent spectral resolution $\mathrm{R}_j(\lambda)$, the instrumental wavelength dependent LSF of the spaxel in angstroms is given by $\sigma(\lambda)_{inst,j} = \frac{\lambda}{f \mathrm{R}_j(\lambda)}$, where $f = \sqrt{8 \ln 2}$ for a Gaussian LSF. The binned LSF is then defined by a nominal propagation of the LSF of each binned spaxel in each spectral pixel. Precisely, under the assumption that each LSF is Gaussian, we define the binned LSF as

the second moment of the distribution defined by the sum of the Gaussian LSFs of each spaxel in bin i:

$$\sigma(\lambda)_{inst,i} = \sqrt{\frac{1}{N_i}\sum_{j=1}^{N_i}\sigma(\lambda)^2_{inst,j}} \tag{3.8}$$

In the case that a single, wavelength-dependent spectral resolution is provided, this equation simplifies such that the binned spectral resolution is simply the provided spectral resolution.

Dependent on the release version analysed, the MaNGA data provides either a single, wavelength dependent resolution channel, defined as the median of the spectral resolutions of all fibres within the IFU (MPL-6 and prior releases), or a 3D resolution cube, providing the resolution for each spaxel (MPL-7 and subsequent releases). The effects of applying the spatially resolved spectral resolutions as opposed to a single median resolution array were explored through Monte-Carlo simulations of stellar kinematic recovery, the methodology and results of which are discussed at in section 3.4.3.

3.4 Template Preparation and Resolution Matching

3.4.1 Template Preparation

The main fitting engine for the stellar-continuum analysis module in SCRAPPy is the penalised pixel-fitting method (`pPXF`), originally developed by Cappellari & Emsellem (2004), later upgraded by Cappellari (2017). The reader is directed to the mentioned papers for a full description of the method. To summarise, `pPXF` assumes a galaxy spectrum can be represented by a number of kinematic components, typically stellar and nebular emission lines, constructed from a set of input templates convolved with their associated component's LOSVD. Currently, SCRAPPy allows only a single kinematic component for the stellar continuum, but allows for up to three distinct emission line kinematic components to represent the systemic gas component, the gas contained within ionised outflows and, in the case of type-1 AGN, an extremely broad set of Balmer series templates to represent the BLR. Within the first two emission line components, the user may also define sub-components to investigate interesting kinematic structures. Whilst some outflows maybe be best represented by more than two Gaussians, further testing and optimisations of SCRAPPy are needed to ensure robustness and accuracy. Finally, SCRAPPy allows for the use of additive and/or multiplicative Legendre polynomials to correct the continuum shape.

SCRAPPy has been designed to allow for input of any spectral template library the user wishes, allowing either empirical or theoretical libraries, consisting of either individual stars

or simple stellar populations. The user must simply ensure the input library is correctly formatted for SCRAPPy, and provide the resolution and wavelength arrays. As `pPXF` requires logarithmically sampled templates, SCRAPPy can also perform logarithmic resampling of templates that are linearly spaced in wavelength, ensuring the data and templates are sampled at the same velocity scale. Furthermore, we ensure the templates and data are both sampled in either vacuum or air wavelengths to avoid disparities, converting the templates to the sampling of the data.

The choice of library is dependent on the properties the user wishes to explore. For chapter 5, we employed the full empirical `MILES` stellar spectral library (Sánchez-Blázquez et al. 2006) , sampled at a spectral resolution of $\Delta\lambda$=2.51Å with a spectral sampling of $\delta\lambda$=0.9Å. These spectra cover the wavelength range 3525Å-7500Å which, although not covering the entire wavelength range afforded by MaNGA, covers the stellar continuum beneath the primary nebular emission lines used to trace outflows, namely [OII]$\lambda\lambda$3727,2730, Hβ, [OIII]$\lambda\lambda$4960,5008, [OI]$\lambda\lambda$6300,6366 [NII]$\lambda\lambda$6550,6585, Hα and [SII]$\lambda\lambda$6718,6733, whilst possessing a spectral resolution high enough to match the MaNGA spectra. In this study, we simply wished to remove the stellar continuum in an accurate yet efficient manner, and utilised the template trimming module (discussed in section 3.4.4) of SCRAPPy to achieve this. The spectra were initially sampled linearly in wavelength space, and thus were logarithmically resampled to the same spectral sampling as the MaNGA data of $\delta\log\lambda=10^{-4}$, which corresponds to a pixel velocity scale of $\Delta V \sim 69$ km s^{-1}.

For our studies in chapter 6, however, we required the stellar populations be recovered, and as such utilised the `MILES` SSP library (Falcón-Barroso et al. 2011). These SSPs cover the same wavelength range and are sampled at the same resolution as the `MILES` stellar templates, and again were logarithmically resampled from their initial linear wavelength grid. We required the templates form a regular gird for the population synthesis module (discussed in section 3.6.7), and as such did not trim the library size.

3.4.2 Resolution Matching

Typically, the provided template spectra will be sampled at a different resolution to the data. As such, a common approach to accurately recover the stellar LOSVD, especially in the Gauss-Hermite parametrisation of the LOSVD (van der Marel & Franx, 1993) which is defined with respect to the Gaussian dispersion, is to match the template resolution to the galaxy data. This resolution matching must be performed with the galaxy data and the templates in the same frame, by either shifting the galaxy data to the rest-frame or redshifting the templates via known galactic heliocentric redshifts, as in the vast majority of cases the spectral resolutions of both will vary with wavelength. Thus, differences between the resolutions will invariably arise

in shifting frame post matching. SCRAPPy, in its data loading and analysis module, shifts all galaxy data, including the resolution, to their rest-frame.

As such, should resolution arrays for both the templates and the data be provided, SCRAPPy will match the template resolution to that of the data by convolving each template spectrum with a wavelength-dependent convolution kernel. This is achieved by convolving each template's flux density in each discretely sampled wavelength channel, $f(\lambda)$ with a Gaussian kernel, $g(\lambda)$, whose width varies with wavelength and is defined in terms of the difference between the spectral resolution of the data and the template. Specifically, for a spectrum and template with wavelength dependent resolutions R_S and R_T respectively, where $R_T > R_S$, the wavelength dependent Gaussian difference between the resolutions in angstroms, $\sigma_{\lambda,\delta}$, can be defined as

$$\sigma_{\lambda,\delta} = \frac{\lambda}{\sqrt{8\ln 2}}\sqrt{\frac{1}{R_{S,\lambda}^2} - \frac{1}{R_{T,\lambda}^2}} \tag{3.9}$$

This, in turn, is used in the general convolution formula to recover the convolved template, $\tau(\lambda)$, defined as

$$\begin{aligned}\tau(\lambda) = (f * g)(\lambda) &= \int_{-\infty}^{\infty} f(\Lambda)\, g(\lambda - \Lambda, \sigma_{\Lambda,\delta})\, d\Lambda \\ &= \int_{-\infty}^{\infty} f(\Lambda)\frac{1}{\sqrt{2\pi}\sigma_{\Lambda,\delta}}\exp\left(-\frac{(\lambda-\Lambda)^2}{2\sigma_{\Lambda,\delta}^2}\right) d\Lambda\end{aligned} \tag{3.10}$$

To always ensure the total flux density within the spectrum remains constant, the kernel here is always normalised to unity. Due to the way in which $\sigma_{\lambda,\delta}$ is defined, the integral over the kernel can be vastly different from unity in regions where the spectral resolution of the template is lower than that of the data, or near to the edges of the spectral range. In these regions, the kernel itself is often no wider than a single spectral channel, and as such is simply replaced by a Kronecker delta function. SCRAPPy also provides the user the choice to either mask these regions, should they be worried the regions will affect the final recovered kinematics and continuum. This will effectively trim the final spectral range analysed by SCRAPPy, and can lead to important stellar features, absorption indices or emission lines not being analysed. For our analyses of the MaNGA datacubes, the resolutions of the MILES templates used were lower for the bluest regions of the spectra (typically around $\lambda \lesssim$4000Å), which contains both the D4000 and D_n4000 stellar indices, as well as the [OII]$\lambda\lambda$3727,3730 emission lines, and as such, we elected to fit the continuum in these areas.

Given the wavelength dependence of the kernel (equation 3.9) one cannot simply perform a convolution using FFTs, and instead a direct summation approach has been adopted. Whilst the computation time for a single, 1D resolution array is negligible, when the user supplies a

3D resolution cube, the template convolution must be performed for each spaxel. However, the additional computation time this introduces to SCRAPPy remains negligible in comparison to the fitting procedure itself, and thus improvements to other modules would have a more significant affect on reducing computation times. Finally, a resolution mask is created to mask the edges of the wavelength range to limit the effect of convolution antialiasing. The width of the edge masks are set by the maximum spectral resolution difference.

3.4.3 Spatially resolved spectral resolution: how much of an improvement?

As SCRAPPy performs matching of the spectral resolution of the input templates to that of the data, any differences between the input data resolution and the "true" data resolution can lead to poor kinematic recovery. An example of this is the transition from the early MaNGA data releases (MPL-6 and prior) data releases to the more recent (MPL-7 onwards) data releases, wherein a 3D extension containing the wavelength dependent LSF for each spaxel was included in place of a 1D median resolution array. This difference is showcased in figure 3.4, where the LSF has been extracted for each spaxel at the observed wavelengths of both Hα and [OIII]λ5008 and is shown in the upper panels. The lower panel shows the median LSF (from the equivalent MPL-5 datacube) in magenta, with the 2.5^{th} and 97.5^{th} percentile LSFs, determined by their quadrature difference from the median LSF, shown as red and blue dashed lines.

We can begin to consider how this variance in spectral resolution across the field of view will affect our recovered kinematics by returning to equation 3.9, alongside the knowledge that the convolution of two normalised Gaussians with dispersions σ_1 and σ_2 is another normalised Gaussian with width $\sigma_3 = (\sigma_1^2 + \sigma_1^2)^{(1/2)}$. Equation 3.9 defined the wavelength dependent width of the normalised Gaussian, $\sigma(\lambda)_\delta$, used as the kernel for convolution of the templates, which we redefine here in terms of the wavelength dependent Gaussian LSF widths of the spectrograph, $\sigma(\lambda)_{inst}$, and the templates, $\sigma(\lambda)_{temp}$, as:

$$\sigma(\lambda)_\delta = \sqrt{\sigma(\lambda)^2_{inst} - \sigma(\lambda)^2_{temp}} \tag{3.11}$$

If we were to assume that the (instrumental) LSF we were to use was either under- or over-estimated, as would be the case if we used the median (magenta) LSF from figure 3.4 in place of the true LSF, then the LOSVD recovered by `pPXF` would be equivalent to a convolution of the real LOSVD with a Gaussian of width given by the quadratic difference between the LSFs. The over-estimation of the dispersion, Δ_σ, can be expressed as:

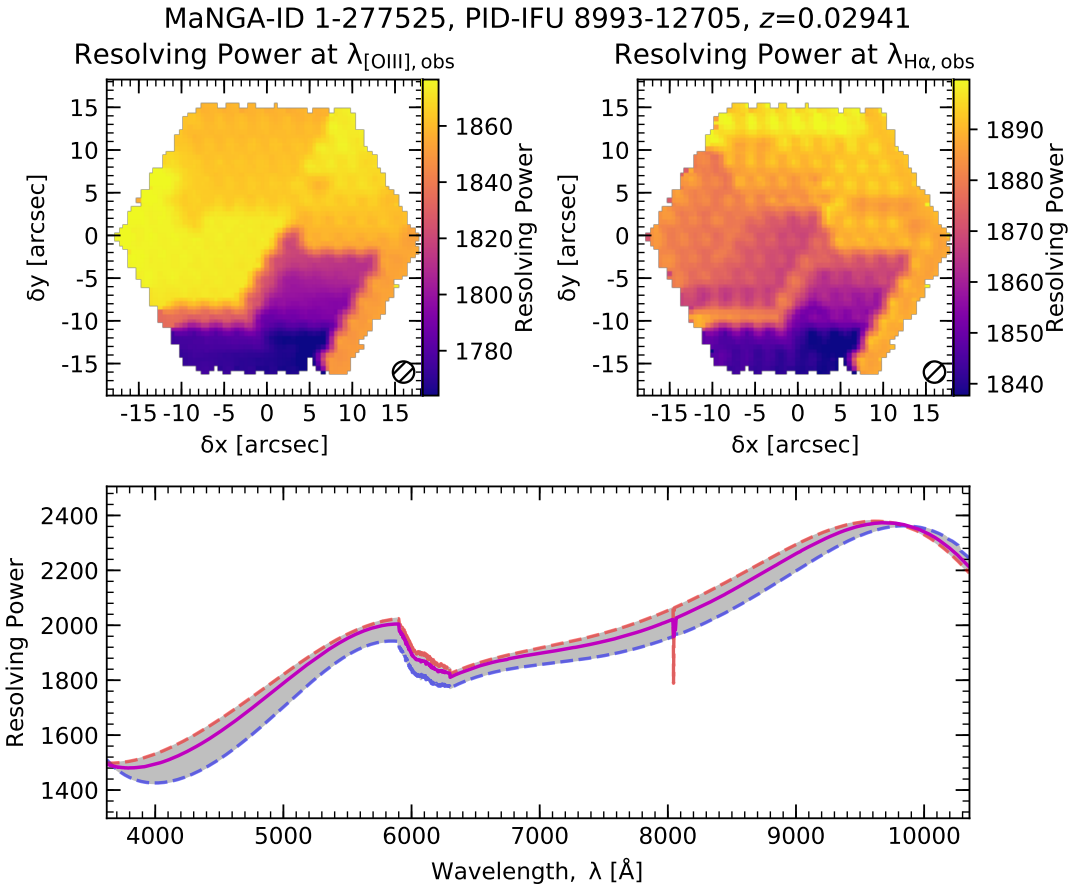


Figure 3.4: An example of the spatially resolved spectral resolutions provided by the MaNGA datacubes. The top left and top right panels show the resolving power at the observed wavelengths of [OIII]λ5008 and Hα, respectively. The bottom panel shows the variation in resolving power across all spaxels for the MPL-7 datacubes, constrained between the 2.5^{th} and 97.5^{th} percentile LSFs, shown as the red and blue dashed lines, respectively. The solid magenta line is the 1D median spectral resolution array provided in the MPL-5 datacubes.

$$\begin{aligned} \Delta_\sigma &= \sigma_{median} - \sigma_{LOSVD} \\ &= \sqrt{\sigma^2_{LOSVD} + \sigma^2_\Delta} - \sigma_{LOSVD} \\ &= \sigma_{LOSVD}\left(\sqrt{1 + \left(\frac{\sigma_\Delta}{\sigma_{LOSVD}}\right)^2} - 1\right) \end{aligned} \tag{3.12}$$

where σ_{median} and σ_{LOSVD} are the velocity dispersions of the fit with templates matched to the median LSF and the true underlying LOSVD of the galaxy, respectively, and σ_Δ is the quadratic difference between the true LSF and the median LSF. As the quadratic difference varies with wavelength, we cannot simply specify how much of an offset will be produced.

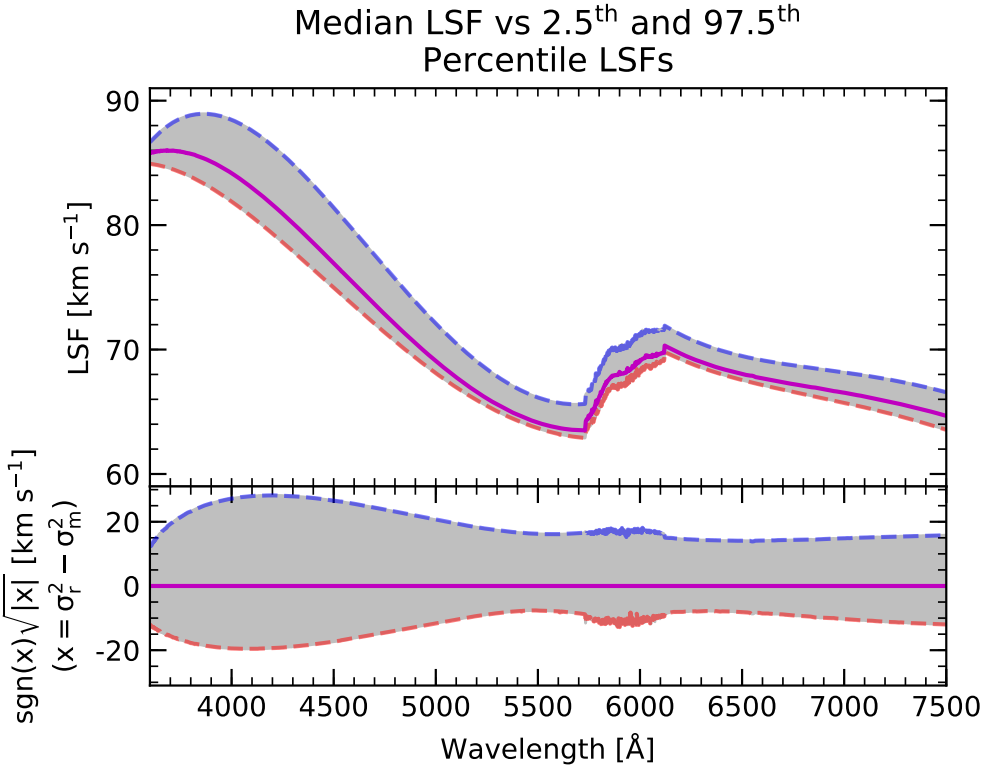


Figure 3.5: Comparisons between the 2.5^{th} and 97.5^{th} percentile LSFs of the MPL-7 datacubes (blue and red dahsed lines respectively), in comparison to the MPL-5 median resolution (magenta line), highlighting the potential offset in the velocity dispersion measurements at each wavelength (in the MILES template wavelength range).

To explore this effect we have run kinematic recovery Monte-Carlo simulations using two LSFs, namely the median LSF and the 2.5^{th} percentile, which we refer to henceforth as our "false" and "true" LSFs respectively. As we wish to be able to control all input for these simulations, we elect to use the `MILES` SSP templates as our basis. We generate a mock galaxy spectrum from the templates by first normalising all templates to a median value of 1. This was done so as to not bias our generated spectra towards the more luminous young SSP spectra, as for the younger SSP spectra, some of the very distinct stellar absorption features, notably the Balmer absorption lines, are more defined. As such, these may have biased our kinematic recovery to perform better in the simulations than we would observe for real galaxy spectra. Taking our normalised spectra, we assign each a random weight between 0 and 1, with the final spectrum being a simple linear stacking of all weighted spectra (e.g. equation 3.5). Following this, we generated a Gauss-Hermite LOSVD Kernel, with a random velocity, velocity dispersion and first two hermite moments h_3 and h_4, constrained to fall within the bounds -300 km s^{-1} < v_{in} < 300 km s^{-1}, 10 km s^{-1} < σ_{in} < 150 km s^{-1}, and -0.1 < h_3 & h_4 < 0.1. The stacked spectrum

was convolved with this LOSVD, then resolution matched to the "true" LSF, as discussed above, Finally, we added random Poissonian noise at an SNR of 30, typical of the central regions of galaxies in the MaNGA sample where we are most likely to detect outflows, defined such that the median flux within the SDSS r-band of our mock spectra divided by the mean of the noise would reach this level. This process was repeated 3,000 times.

We then generated two libraries from the default MILES templates: a first template library resolution matched to the "false" median LSF of the MPL-5 data, and a second resolution matched to the "true" 2.5^{th} percentile LSF. We then used `pPXF` to recover the stellar LOSVD, and all output components were saved.

Figure 3.6 shows the results of these simulations. The panels from top to bottom show the velocity difference, $v_{in} - v_{out}$, velocity dispersion difference, $\sigma_{in} - \sigma_{out}$, and the quadratic difference between the velocity dispersions. In each panel, the median offset and standard deviation, in bins of width 10 km s^{-1}, are shown as the solid black and dashed black lines respectively. Whilst we see no improvement in the recovery of the velocity, we see that the velocity dispersion is significantly better recovered when a full spaxel resolution matching (left, middle panel) is applied as opposed to a median resolution matching (right, middle panel). We can further compare these results by considering the difference between the median trends for $\sigma_{median,\ out}$ and $\sigma_{spaxel,\ out}$, with respect to equation 3.12. This equation is plotted as curves in figure 3.7 for varying single (non-wavelength dependent) values of the quadratic difference between the median LSF and the "true" underlying LSF, σ_{Δ}. The true σ_{Δ} between the median and resolution matched median trends is shown as the magenta line, with the standard deviation shown as magenta dashed lines. The observed offset agrees with the theoretically predicted offset curves, suggesting the offset may be between 25 km s^{-1} and 29 km s^{-1}. Figure 3.5 shows the quadratic difference between the median LSF and the 2.5^{th} and 97.5^{th} percentile LSFs, again as blue and red dashed curves respectively. This plot shows that the offset between the LSFs is strongest between approximately 4000Å and 4500Å, in a region dominated by strong stellar features. This suggests that the observed dispersion differences may be driven by kinematics being determined by the strong features that fall in this region.

These results indicate that ensuring the resolution of the templates is matched to the correct LSF of the galaxy can lead to improvements in the recovery of the kinematics of the stellar continuum. Furthermore, the offsets observed here due to the differences between the 'true' LSF and an assumed LSF will propagate into recovery of the LOSVD of the emission lines, affecting flux measurements. Hence, for our analyses of the MaNGA data in chapter 6 (for which we use the MPL-7 data that provides spatially resolved LSFs) we preform template resolution matching for each spectrum.

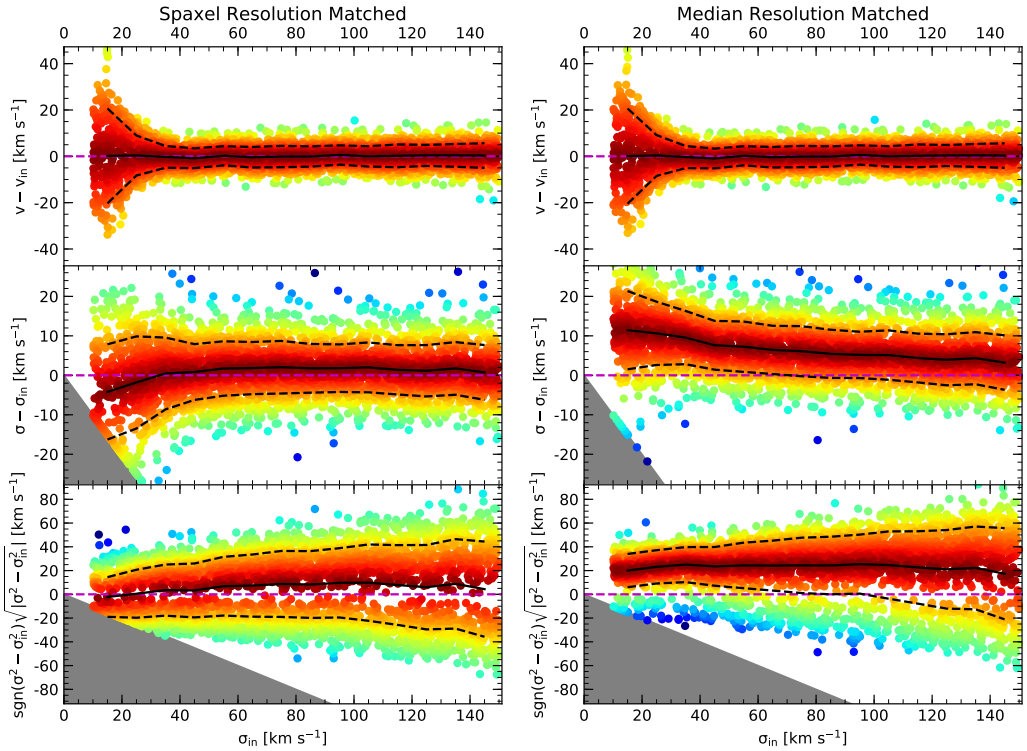


Figure 3.6: Comparisons of the kinematic recovery from Monte-Carlo simulations of galaxy spectra. The left panels show the recovered kinematics in comparison to the input kinematics for templates matched to the spectral resolution of the simulated galactic spectrum, with the resolution chosen as the 2.5^{th} percentile LSF of a MaNGA MPL-7 datacube, as in figures 3.4 and 3.8. The right panels show the same but for templates matched to the MPL-5 median resolution of the same galaxy, not to the resolution of the generated spectrum. The 2.5^{th} percentile resolution was chosen to explore the maximum possible offsets between recovered values and input values for incorrectly resolution matched templates. Whilst there appears to be little to no affect on the recovery of the velocity, the dispersion recovery is significantly affected, with an offset of ~10% (~7 km s^{-1}) at 69 km s^{-1}, the velocity scale of the MaNGA data. The dip observed at low σ in both of the middle panels is due to the `MILES` templates used being undersampled. The dispersions here are lower than the widths of the spectral channels, with the dip beginning around 35 km s^{-1}, equivalent to half the pixel velocity scale of the `MILES` templates.

3.4.4 Template library trimmming

As the execution time for the `pPXF` algorithm is typically larger than $O(N_{temp})$ for a set of N_{temp} templates (both stellar and emission line component templates), SCRAPPy offers some functionality to reduce computation time. This is achieved by selecting a representative subsample of stellar templates from the complete stellar library. Whilst computation time is typically less of an issue with single objects, for large surveys like MaNGA, or even for large

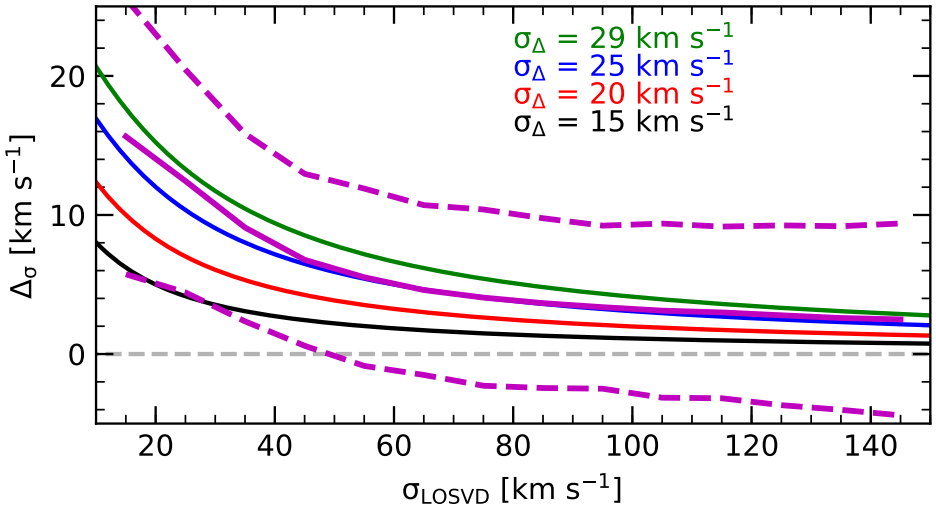


Figure 3.7: Comparison of the overestimation of the dispersion as a function of input velocity dispersion. The black, red, blue and green lines show the predicted overestimation for different maximum quadratic differences between the "true" and median LSFs, calculated using equation 3.12, of 15 km s^{-1}, 20 km s^{-1}, 25 km s^{-1} and 29 km s^{-1}, respectively. The magenta line shows the observed overestimation, showing that the value falls between 25 km s^{-1} and 29 km s^{-1}, indicating that correct template resolution matching is important for good kinematic recovery.

datacubes from new IFUs (e.g. MUSE, with a field of view of $1'\times1'$, and a spatial resolution of $0.2''$, resulting in ~90,000 spectra), an optimal selection of templates can be critical to meeting computational needs. SCRAPPy selects this minimal template selection thorough the fitting the spectra of a number of user defined regions within each galaxy, and/or from the total galaxy integrated spectrum from all good spaxels (defined by the binning module), extracting all templates with non-zero weightings from the `pPXF` fit matrix.

For chapter 5, this trimming was based upon a fit to the integrated spectrum using the full 985 templates of the `MILES` stellar library, selecting any non-zero weighted templates as the `MILES-trim` library for each galaxy. Typically, this reduced the full library to a subsample of between 25 and 40 templates. Other methodologies for creating optimal template selections have been previously employed, such as for the MaNGA DAP where a sub-sample was chosen to select stars that uniformly sampled a grid in known stellar physical parameters, such as temperature (T_{eff}), metallicity ([Fe/H]), and surface gravity (g) (e.g. Shetty & Cappellari 2015). Whilst these methods certainly decrease the computational needs, and are appropriate for determination of stellar kinematics, some deficiencies in the trimmed libraries for general continuum fitting have been previsouly observed (Belfiore et al., 2019).

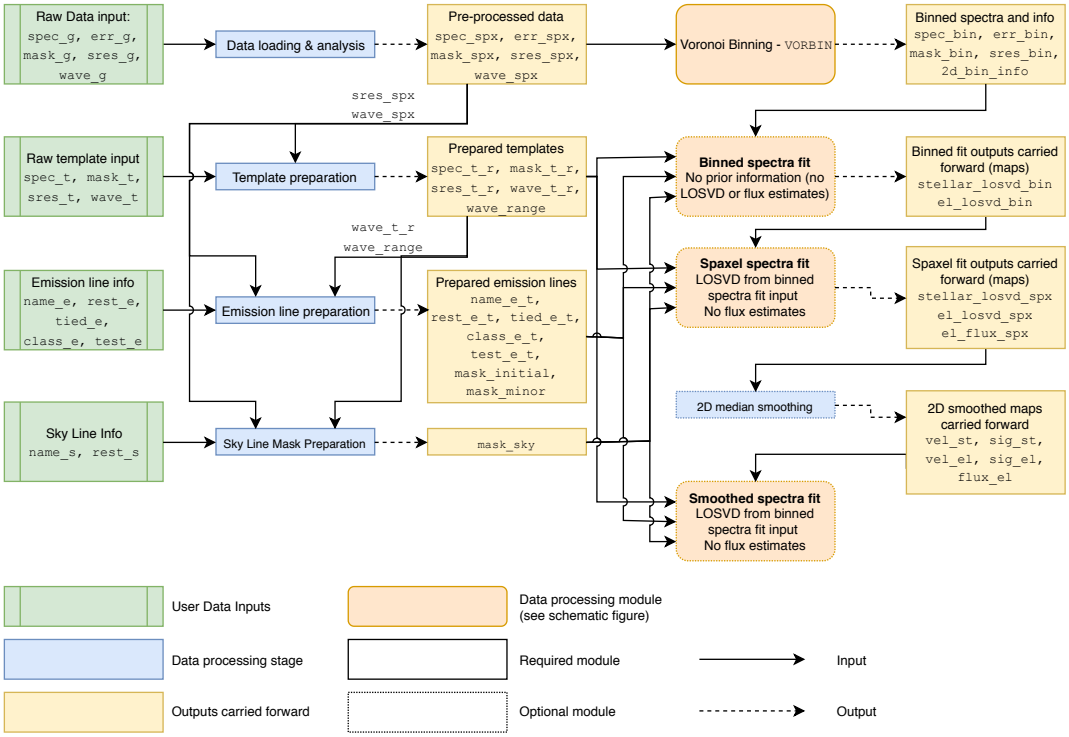


Figure 3.8: Schematic of the work flow of the SCRAPPy pipeline, showing how inputs and outputs from loading modules, preparation modules and analysis modules feed into each other. The data processing modules are described in section 3.6, with an algorithmic flow diagram shown in figure 3.9

3.5 Fitting stages of SCRAPPy

SCRAPPy performs fits to the data in an iterative manner, using recovered fluxes and kinematics from previous stages to improve subsequent fits. The full spectral and emission line fitter around which most of the stages operate is discussed later in section 3.6. Here, we detail the calling sequence of the main SCRAPPy wrapper module. The final call sequence contains optional modules that the user may choose to include, however here we will discuss the full call sequence, denoting optional modules.

SCRAPPy operates in the following stages:

1) **Fit the spectra of user defined areas (e.g. central regions) to determine the number of required components.** The purpose of this stage is to ensure SCRAPPy knows the maximum number of components needed for subsequent fits. As outflows are typically nuclear, being driven by AGN or centrally enhanced SF, SCRAPPy can automatically define a central spectrum by co-adding the spectra located within an annulus, of radius defined by the user but recommended as the instrumental PSF radius, around the highest

continuum flux spaxel. This is used if no other region is given by the user. The fits to these regions define the maximum number of components used throughout all other fits (i.e. 1 or 2 components fit, and type-1 AGN presence).

2) **OPTIONAL: Fit the total integrated spectrum of the galaxy to determine a trimmed template library.** This stage serves to reduce the template library size to decrease computational time. The total integrated spectrum is computed from all good spaxels defined by the binning module, coadded using the same formulae as for the binned specta. This trimmed library is only used for stellar continuum and kinematics recovery, and is not used for population synthesis (section 3.6.7) as this requires a regular grid of template properties. This stage is entirely optional.
3) **Fit all binned spectra.** This is the minimum amount of data processing and analysis that SCRAPPy will undertake without the user requesting further stages be performed. Each binned spectrum fit is independent of all other spectra, and as such this stage can be processed using multiple computational cores to speed up processing. Population synthesis is only performed on binned spectra (if requested) as fits to spaxels with low SNR will be noisy and inaccurate. This stage produces two datacubes: the first contains all of the input data and output fits from `pPXF` and out emission line fitter; the second contains 2D maps of the kinematics and fluxes of both the stellar and emission line components. Population maps are not produced, but the weighting of the templates in the final population synthesis fit are saved, as different definitions of properties (e.g. age or metallicity) may be required by the user.
4) **OPTIONAL: Use information from binned spectra to fit spectra of each bin's constituent spaxels.** The spectra of spaxels that are constituent to a bin are refitted, using the kinematics of the binned fit as initial LOSVD estimates for both the stellar and emission line components. Again, two datacubes are produced containing the fittings and 2D property maps.
5) **OPTIONAL: Perform 2D median filter smoothing on recovered property maps (v, σ, flux) and use these as initial parameter estimates in subsequent spaxel spectra refits.** Taking the output 2D maps of the spaxel spectra fittings, SCRAPPy performs a 2D median filter smoothing of the maps (discussed in section 3.7.1), with these providing initial LOSVD and flux estimates for refits to the spaxel spectra.

3.6 Full Spectral Fitting and Emission Line Analysis

The main module around which SCRAPPy operates is the full spectral and emission line analysis module. Due to the iterative nature of SCRAPPy, this module is called multiple times

in a single complete SCRAPPy run-through, with each call using data from previous fits. Figure 3.9 shows the flow of this module, with each stage detailed throughout this section.

3.6.1 Mask creation

The first stage in the full spectral fitting module of SCRAPPy aims to recover the stellar LOSVD and an initial continuum estimate. Whilst `pPXF` has been shown to be sufficient in simultaneously modelling the stellar continuum and emission lines, it does so by treating them in the same manner: both are provided as templates (stellar templates and Gaussian templates of a given initial width) that are convolved with a Gaussian (or Gauss-Hermite) LOSVD kernel. Whilst in new versions of pPXF this has been accounted for through performing fits in Fourier space, this pipeline currently uses an historic version of pPXF and, as such, this methodology can lead to a "pixelation" of the emission lines, especially for narrow lines with widths comparable or less than the instrumental dispersion. Furthermore, if an incorrect number of kinematic components is assumed, issues may arise with recovery of the stellar LOSVD and subsequent continuum extraction. Hence, as we do not wish to analyse emission lines at this point, an emission line library is only used at this stage to construct an emission line mask. Additionally, a sky line library may also be included to generate a sky mask. By default, SCRAPPy contains an emission line table containing nebular emission and absorption lines within the nominal wavelength range 3200Å< λ_{rest} < 9600Å, and a sky line mask containing a number of bright Oxygen night-sky lines within the optical wavelength range. Most science ready data will have had sky subtraction performed but residuals of these bright night-sky lines are almost ubiquitous, hence SCRAPPy allows for the masking of these lines. Tables 3.1 and 3.2 detail the lines contained within each library respectively.

SCRAPPy actually constructs two emission line masks, defined by the lines the user wishes to use to explore the gas kinematics. In the emission line input file, the user may define "major" and "minor" lines to be fit: "major" lines are those the user wishes SCRAPPy to use in the determination and fitting of potential multiple components, whilst "minor" lines are those expected to be present in the galactic data, but are typically faint and thus may affect kinematics recovery, whilst also increasing computation time. These two masks are deployed simultaneously for the initial stellar LOSVD and continuum recovery, but the "major" line mask is discarded earlier in the fitting procedures than the "minor" line mask.

SCRAPPy constructs the emission line masks by finding all channels within a velocity range set by the user, or by default within ±1000 km s^{-1}, centred on the rest wavelength of each emission line provided. Any pixels that fall fully or partially within this range are masked. Whilst this wavelength range may seem large, the main purpose of SCRAPPy is for the detection and analysis of outflows, which may exhibit velocities in the few ∼100 km s^{-1}

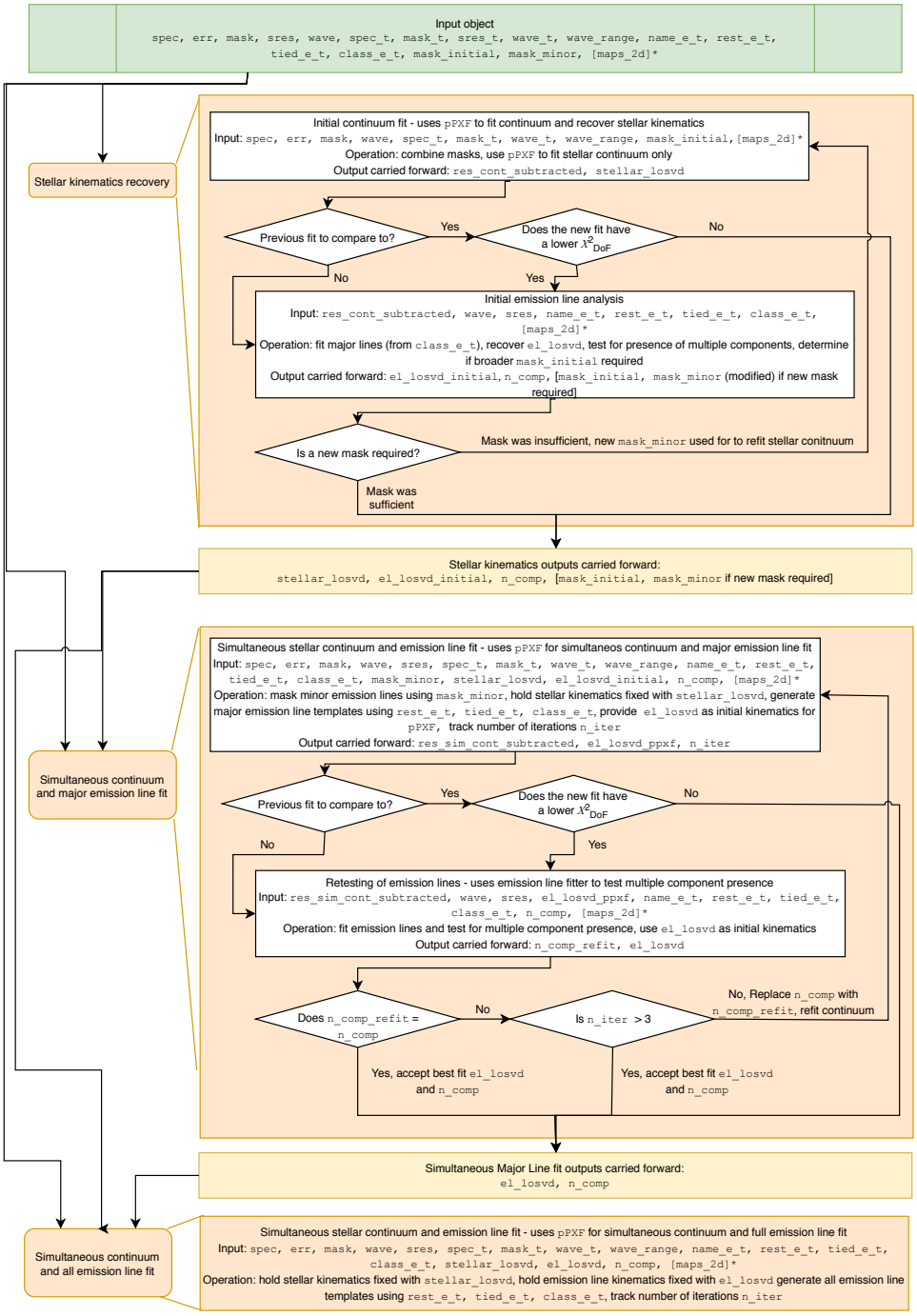


Figure 3.9: Algorithmic flow diagram of the full spectral fitting and emission line analysis module, detailed in section 3.6. Order of progression is from top to bottom, with inputs into modules shown as arrows. The orange boxes show the flow within each sub-module, with relevant outputs to be carried to the next submodule shown within the yellow boxes. The green input object is that produced by the initial SCRAPPy pre-processing pipeline, shown in figure 3.8

ID	Ion	λ_{rest}	Fit/Ignore	Major/Minor	Line/Doublet	Relative Intensity	Intensity Limited	Intensity Limiting Line	Intensity Limits
0	HeII3204	3204.038	f	minor	1	1	0	0	0
1	NeV3347	3346.783	f	minor	1	1	0	0	0
2	NeV3427	3426.863	f	minor	1	1	0	0	0
3	H23	3670.5154	i	minor	1	1	0	0	0
4	H22	3672.5279	i	minor	1	1	0	0	0
5	H21	3674.8109	i	minor	1	1	0	0	0
6	H20	3677.416	i	minor	1	1	0	0	0
7	H19	3680.4065	i	minor	1	1	0	0	0
8	H18	3683.8627	i	minor	1	1	0	0	0
9	H17	3687.887	i	minor	1	1	0	0	0
10	H16	3692.6119	i	minor	1	1	0	0	0
11	H15	3698.2104	i	minor	1	1	0	0	0
12	H14	3704.9133	i	minor	1	1	0	0	0
13	H13	3713.0334	i	minor	1	1	0	0	0
14	H12	3723.0035	i	minor	1	1	0	0	0
15	OII3727	3727.092	f	minor	1	1	0	1	0
16	OII3730	3729.875	f	minor	1	1	1	0	0.35,1.5
17	H11	3735.4365	i	minor	1	1	0	0	0
18	H10	3751.2243	i	minor	1	1	0	0	0
19	H9	3771.708	i	minor	1	1	0	0	0
20	H8	3798.9826	i	minor	1	1	0	0	0
21	H7	3836.479	i	minor	1	1	0	0	0
22	NeIII3870	3869.86	f	minor	1	1	0	0	0
23	H6	3890.1576	f	minor	1	1	0	0	0
24	NeIII3970	3968.59	f	minor	1	1	0	0	0
25	H5	3971.202	f	minor	1	1	0	0	0
26	SII4069	4068.6	f	minor	1	1	0	0	0
27	SII4076	4076.349	f	minor	1	1	0	0	0
28	Hd	4102.8991	f	minor	1	1	0	0	0
29	Hg	4341.691	f	minor	1	1	0	0	0
30	OIII4364	4364.435	f	minor	1	1	0	0	0
31	HeII4687	4687.015	f	minor	1	1	0	0	0
32	ArIV4713	4712.58	f	minor	1	1	0	0	0
33	ArIV4742	4741.45	f	minor	1	1	0	0	0
34	Hb	4862.691	f	major	1	1	0	0	0
35	OIII4960	4960.295	f	major	d36	0.35	0	0	0
36	OIII5008	5008.24	f	major	1	1	0	0	0
37	NI5199	5199.349	f	minor	1	1	0	0	0
38	NI5202	5201.705	f	minor	1	1	0	0	0
39	HeI5877	5877.243	f	minor	1	1	0	0	0
40	NaI5892	5891.583	i	minor	1	-1	0	0	0
41	NaI5898	5897.558	i	minor	1	-1	0	0	0
42	OI6302	6302.046	f	minor	1	1	0	0	0
43	OI6366	6365.535	f	minor	d42	0.333	0	0	0
44	NII6550	6549.86	f	major	d46	0.34	0	0	0
45	Ha	6564.632	f	major	1	1	0	0	0
46	NII6585	6585.271	f	major	1	1	0	0	0
47	SII6718	6718.294	f	major	1	1	2	0	0.44,1.5
48	SII6733	6732.674	f	major	1	1	0	2	0
49	ArIII7138	7137.76	f	minor	1	1	0	0	0
50	ArIV7173	7172.68	f	minor	1	1	0	0	0
51	ArIV7240	7239.76	f	minor	1	1	0	0	0
52	ArIV7265	7265.33	f	minor	1	1	0	0	0
53	OII7321	7320.94	f	minor	1	1	0	0	0
54	OII7322	7322.01	f	minor	1	1	0	0	0
55	OII7332	7331.69	f	minor	1	1	0	0	0
56	OII7333	7332.75	f	minor	1	1	0	0	0
57	ArIV7334	7334.17	f	minor	1	1	0	0	0
58	ArIII7753	7753.24	f	minor	1	1	0	0	0
59	ArIII8039	8038.73	f	minor	1	1	0	0	0
60	SIII8832	8831.8	f	minor	1	1	0	0	0
61	SIII9071	9071.1	f	minor	1	1	0	0	0
62	SIII9533	9533.2	f	minor	1	1	0	0	0

Table 3.1: Default nebular emission and absorption line library used by SCRAPPy. The columns provide the following information: 1) Line ID; 2) Ion Name, as used in SCRAPPy output files; 3) Rest wavelength, provided as Ritz wavelengths in vacuum from the National Institute of Standards and Technology (NIST); 4) Fit (f) or Ignore (i) indicators; 5) Major or Minor line identifier; 6) Line/Doublet identifier. d*n* identifies a doublet line, *n* indicates the ID of the line to which the intensity should be tied; 7) Relative Intensity of the line (used as initial amplitude estimates), 1 for emission lines, -1 for absorption lines. For tied doublets, this identifies the relative intensity to the main doublet line. 8) Intensity limiting identifier, defines the line for which the intensity will be limited; 9) Intensity limiting line, the main line from which the density limited line intensity will be determined; 10) Intensity limits, (minimum, maximum).

ID	$\lambda_{\rm rest}$
0	5577.4
1	5889.9
2	6306.6
3	6363.7
4	7246.0

Table 3.2: Default sky line masking library used by SCRAPPy. The columns provide the following information: 1) Sky line ID; 2) Rest wavelength, provided as Ritz wavelengths in vacuum

range, with velocity dispersions of a similar magnitude, thus may even still exceed the masks, affecting the stellar continuum fit. A similar prescription is used for the sky lines, except a default width of only 200 km s^{-1} is assumed, with the masks centred on their blueshifted (using the galactic heliocentric redshift provided by the user) central wavelengths.

Additionally, the wavelength ranges of the galaxy data and stellar templates are often not the same. As such, some trimming must be performed, of either the templates or the galactic data dependent on which possesses a greater wavelength coverage. For example, we have tested SCRAPPy extensively using the MaNGA data, which covers the wavelength range ~3500-10500 Å, with both the MILES stellar and SSP spectral libraries, which cover the wavelength range ~3600-7500 Å. As such, the (rest-framed) data was trimmed to cover the same range as the templates.

3.6.2 Initial Stellar Kinematics Analysis

Having processed the data, templates and constructing the required masks, SCRAPPy proceeds to an initial stellar continuum fit. A number of additional key values are required by `pPXF` for the fitting procedure, and can be modified in the input files by the user. Firstly, initial estimates for the stellar LOSVD components to be fit must be provided. By default SCRAPPy uses a Gauss-Hermite LOSVD with 4 moments, providing initial stellar velocity, velocity dispersion, h_3 and h_4 values of 0 km s^{-1}, 100 km s^{-1}, 0 and 0 respectively. `pPXF` can also make use of additive Legendre polynomials, and one of either a reddening curve or multiplicative Legendre polynomials to correct the shape of the fits. The user can disable any of these, though by default only an 8th order additive polynomial is applied. These polynomials are used by `pPXF` to compensate for minor mismatches between the templates and galactic spectra that vary with wavelength which may arise due to flux calibration issues, reddening (both internal and/or foreground) or template library incompleteness . Additive polynomials can affect measurements of weak metal lines that directly influence stellar kinematics and weak emission line wings tracing outflows, and thus can affect the recovered LOSVD. Further, dispersion (and

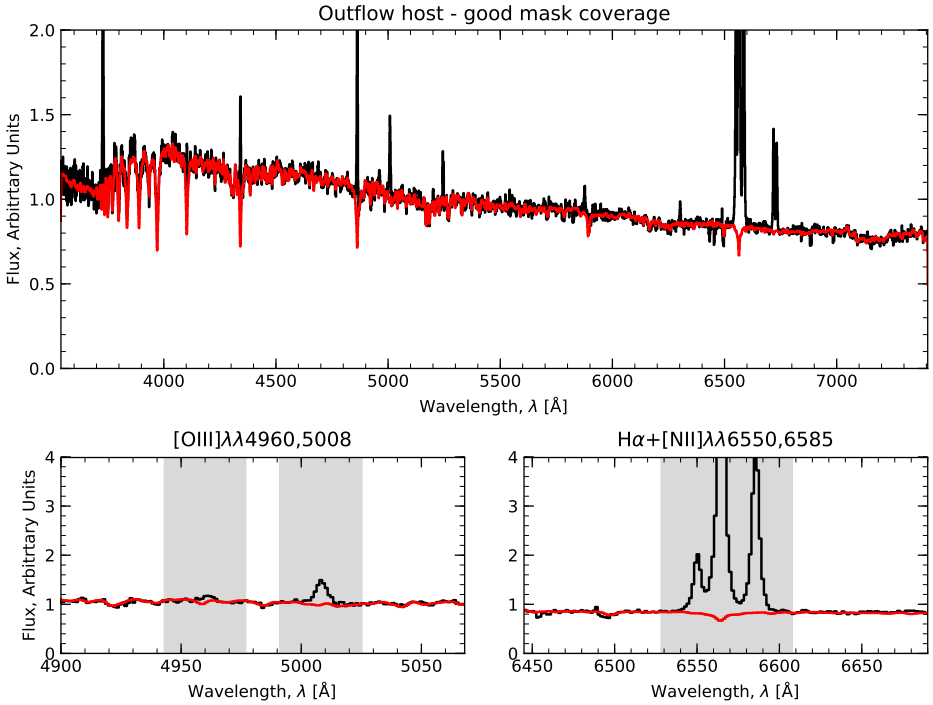
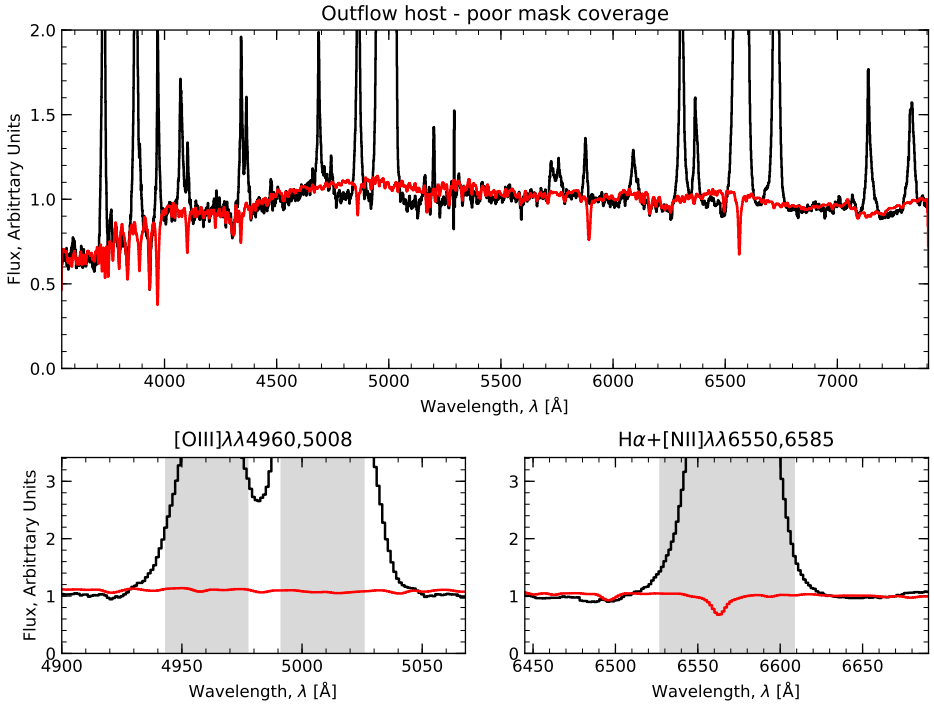


Figure 3.10: A good spectral fitting to an outflow host galaxy. The broad wings of the emission lines are sufficiently masked, thus allowing for a clear recovery of the stellar continuum. The input spectral data in shown in black, whilst the continuum recovered by `pPXF` is shown in red. The grey channels are the emission line masks.

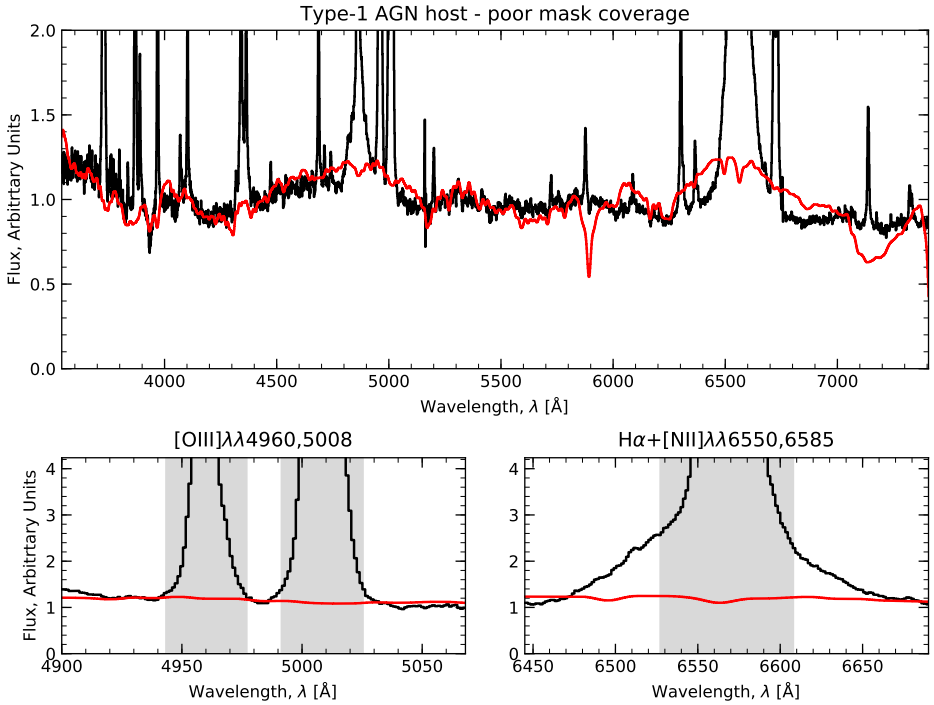
higher moments) can be affected by "template mismatch", i.e. detailed template changes (van der Marel et al., 1994). The impact including polynomials in the fitting has on the derived kinematics, especially the dispersion, has been previously explored, such as in Westfall et al. (2019) where the impact of polynomial order on the stellar dispersion measurements was tested for the MaNGA data. They found that trends in the observed stellar dispersion measured existed for polynomials of order lower than 3, with some trends arising for polynomials of order 9 or greater. They conclude that the systematic change in stellar dispersion as a function of polynomial order is small for orders between 3 and 9, non inclusive.

The aim of this module is not to recover the full stellar continuum, but to both recover the stellar kinematics from stellar features not contaminated by emission lines and to provide a sufficient initial continuum subtraction for analysis of the emission lines. Typically, we find that this module provides good recovery of the stellar kinematics, provided the emission line masks sufficiently mask the broad emission line wings typical of outflows.

Examples of this initial fitting are shown in figures 3.10 and 3.11 for 3 of the galaxies in MaNGA. In figure 3.10 we show an example of this initial fitting module producing a good fit



(a)



(b)

Figure 3.11: (a): A poor initial attempted recovery of the stellar continuum due to contamination from the extremely broad wings of the emission lines tracing outflows. The fit is particularly poor around the [OIII]λ5008 line. The input spectrum is shown in black, the recovered continuum is shown in red and the initial (insufficiently broad) emission line masks are shown as the grey channels. (b): The same as a but for a type-1 AGN host. The extremely broad Blamer lines heavily contaminate the continuum, leading to very poor continuum recovery.

to the stellar continuum, and thus assumed good stellar kinematics, due to sufficient masking of the emission line. On the other hand, figure 3.11 shows two examples where this fit fails. In the top panels, a galaxy with extremely broad emission line wings is shown. The extension of the broad wings past the initial mask forces `pPXF` to minimise the pixel offset by increasing the level of continuum in there regions, of particular notice around the [OIII]λ5008 line. In a similar vein, the bottom panels show a MaNGA galaxy with emission wings far exceeding the initial masks. However, this galaxy is a type-1 AGN host, as evidenced by the extremely broad (FWHM $\sim$10000 km s^{-1}) Balmer emission line profiles (especially noticeable for Hα and Hβ), typical of the BLR. Interestingly, this galaxy also exhibits broad wings in the emission line profiles of other nebular lines, such as [OIII]. In both cases, it is apparent that the initial continuum fit has failed, and the masks must be reconstructed to sufficiently mask the continuum regions contaminated by these lines. These new widths are estimated from an initial emission line fitting. As this continuum may be wrong (such as in the two discussed examples) this initial emission line fit is used to simply recover good initial estimates of all required emission line component kinematics for subsequent fitting module.

3.6.3 Initial Assessment of Emission Lines

The next stage in the module attempts to determine the presence of multiple emission line components, and to assess the extent to which they may exceed the initial masks. Specifically, SCRAPPy generates a residual spectrum for the fit using the initial stellar continuum. This is used to investigate the presence of multiple components in the emission lines by modelling the emission lines as Gaussian profiles, assuming a narrow component tracing the systemic nebular gas in the host galaxy disk (namely virialised motions) and a broad component tracing gas entrained in an outflow. Additionally, this module assesses the presence of extremely broad Balmer line profiles to determine type-1 AGN presence, if requested by the user. This is achieved through a simple chi-squared minimisation using a Levenberg-Marquardt methodology, implemented in python as the `mpfit` package. To ensure efficient computational time, we assess only lines flagged as "major" in the user-input emission line file. By default, SCRAPPy uses 5 major emission lines in this module: Hα, Hβ, and the [OIII]$\lambda\lambda$4960,5008, [NII]$\lambda\lambda$6550,6585 and [SII]$\lambda\lambda$6718,6733 doublets, though in chapters 4 and 5 we also make use of the faint [OI]$\lambda\lambda$6302,6366 doublet. The fitting procedure of the emission lines is discussed below in section 3.6.3.1, followed by a discussion into the assessment of the emission line fits and the process via which we determine the number of kinematic components present in section 3.6.3.2.

3.6.3.1 Emission line fitting

All emission lines are modelled with a single Gaussian profile per kinematic component, imposing the same central velocity and velocity dispersion for all nebular lines within each component. The intensities of the Gaussian components of the individual nebular lines are free to vary, though the user may elect to confine the emission line ratios of certain doublets to their Einstein coefficients (e.g. the [OIII] and [NII] doublets), or for the ratio to vary between two set limits to trace properties such as electron density (e.g. the [SII] doublet). The implementation of the emission line ratios is discussed in detail in section 3.7.2.

As we wish to tie the velocity dispersions of all lines within a kinematic component, SCRAPPy makes a number of data conscious and physically motivated considerations. Firstly, the spectral resolution may vary with wavelength across the spectrum. Assuming the user has provided a spectral resolution that is well characterised in all spectral regions we can link the intrinsic velocity dispersions (widths) of the lines to the observed velocity dispersion through the relation:

$$\sigma_{line} = \sqrt{\sigma_{SRES}^2 + \sigma_{int}^2} \tag{3.13}$$

where σ_{line} is the dispersion of the final Gaussian fit, σ_{SRES} is the spectral resolution at the observed spectral wavelength of the line, and σ_{int} is the tied velocity dispersion between all lines, fit here. If no instrumental resolution has been provided, σ_{SRES} is simply set to 0 km s^{-1}.

Secondly, SCRAPPy assumes that in all cases the second (broad) kinematic component will *always* be broader than the first (narrow) kinematic component. If this is not considered, a swapping of the narrow and broad components, in terms of their locations within the output matrix, can arise, an issue we colloquially refer to as "component swapping". As such, this problem is tackled by SCRAPPy by implementing a restriction on the width of the broad component fit, such that the intrinsic (not broadened by the instrumental LSF) dispersion of the broad component, $\sigma_{int,b}$, is newly defined as:

$$\sigma_{int,b} = \sigma_{int,n} + \sigma_{excess} \tag{3.14}$$

where $\sigma_{int,b}$ is the intrinsic dispersion of the narrow component and σ_{excess} is the excess width of the broad component with respect to the narrow component. This excess can be constrained by the user, and by default cannot fall below 50 km s^{-1}, ensuring the broad component is always broader than the narrow component by this minimum amount. With this definition, the observed velocity dispersion of the narrow component remains given by equation 3.13, but the observed broad component dispersion is instead given by:

$$\begin{aligned}\sigma_{line,b} &= \sqrt{\sigma^2_{int,b} + \sigma^2_{SRES}} \\ &= \sqrt{(\sigma_{int,n} + \sigma_{excess})^2 + \sigma^2_{SRES}}\end{aligned} \tag{3.15}$$

where again if no spectral resolution has been provided by the user σ_{SRES} is set to 0 km s^{-1}.

As the fits are computed using a chi-squared minimisation, spectral windows around the emission lines must be defined. By default, SCRAPPy will select windows ±2000 km s^{-1} from the rest wavelengths of the investigated emission lines, though the user can increase the size of this window. If the wings of the broad components are determined to exceed this window, later calls of the module will extend the spectral range.

3.6.3.2 Determination of kinematic component presence

Depending on the number of different components the user wishes to investigate, SCRAPPy will perform up to four emission line fits (narrow only, narrow and broad, narrow and type-1 AGN, and narrow, broad and type-1 AGN components fits). In the case of including the type-1 AGN component, SCRAPPy first aims to determine whether this component is needed. This component is analysed on the strongest Balmer line available, typically Hα. SCRAPPy requires that both the emission line's SNR (determined from the peak intensity divided by the median error within the FWHM of the line's observed central wavelength) and equivalent width (EW) exceed user set minimum values, set to 3 and 3Å respectively by default. If the type-1 AGN component fails to meet either of these requirements, then the component is removed and a flag set to inform later stages to not fit this component. Similarly, a flag will be set to classify the galaxy as a type-1 AGN host if the component does exceed these limits.

For the remaining kinematic components, SCRAPPy only considers an outflow to be present if the broad emission line component can be clearly distinguished from the narrow nebular emission lines tracing systemic motions of the galaxy. We quantify this by defining a a number of selection criteria imposed on the broad component fits to ensure that, even if a broad component is traced, it is not simply an artefact of poor stellar continuum subtraction or flux calibration errors. SCRAPPy looks at the fits to a user defined subset of the "major" emission lines., henceforth called the "testing" lines. For our analyses of the MaNGA data in searching for ionised outflows, we restricted the initial search to both the Hα and [OIII] lines. For a component to be accepted, the lines must meet the following criteria:

1) **The Signal-to-Noise Ratio (SNR) of the fit broad components testing lines are greater than the set limit (default of 3).** This requirement is driven by previous investigations into SNR levels required to ensure that emission line fits are more significant than any

fitting artefact, whilst the recovered kinematics and line fluxes are also well constrained (Belfiore et al. 2016). With at least one of the testing lines being fit at the required level, the constraints that fit will provide to the kinematics significantly help with recovery of line fluxes at lower SNRs.

2) **The Equivalent Width (EW) of the fit broad components testing lines are greater than the set widths (default of 3Å).** This constraint is adopted to again ensure that the fitting is not simply an artefact of poor stellar continuum subtraction or due to flux calibration issues.
3) **The kinematics of the broad component fits have not hit any of the boundary restrictions placed on the fit.** This, in general, ensures the broad component is always *at least* (by default) 50 km s^{-1} broader in terms of velocity dispersion (equivalent to a FWHM = 120 km s^{-1}) than the narrow component fit to the nebular emission lines, ensuring the component is significant. However, our imposed upper limit also ensures that extremely broad components (σ_v greater than 1000 km s^{-1} and not associated to a type-1 AGN) that may be artefacts of poor stellar subtraction or correcting for noise are removed from out analysis. Further, restrictions to the maximum absolute velocity shift from the emission line rest frame ensure we detect velocities consistent with previous observations of outflows, and do not fit either background artefacts or noise spikes outside of the spectral windows.
4) **The amplitude ratios of the broad component testing lines in comparison the their respective narrow component amplitudes are, at minimum, ≥5%.** We place no restriction on the upper limit for the ratios as some of the galaxies hosting more extreme outflows previously investigated exhibit broad components more luminous than the narrow component.

If none of the testing lines satisfy all four requirements, the broad component fit is immediately rejected. However, should at least one testing line pass all of these initial tests, the emission line fitter will move to the next testing stage. To fully accept the broad component, SCRAPPy requires that the multiple components fit improves upon the reduced chi-squared, χ^2_{DoF}, of the single component fit by a minimum user defined percentage value, by default 10% (i.e. $\chi^2_{DoF,2} < 0.9\chi^2_{DoF,1}$). For the default value of 10%, this methodology is very conservative, as with so many spectral points this always translates into a statistical significance of the additional component higher than 99%. Furthermore, as SCRAPPy performs this fit on a large number of spectra (up to a spaxel-by-spaxel basis) for a datacube, we need a comparison technique that is computationally efficient. This methodology requires no further calculations to determine the best fit to the data, as these values are readily available from the component fitting, and

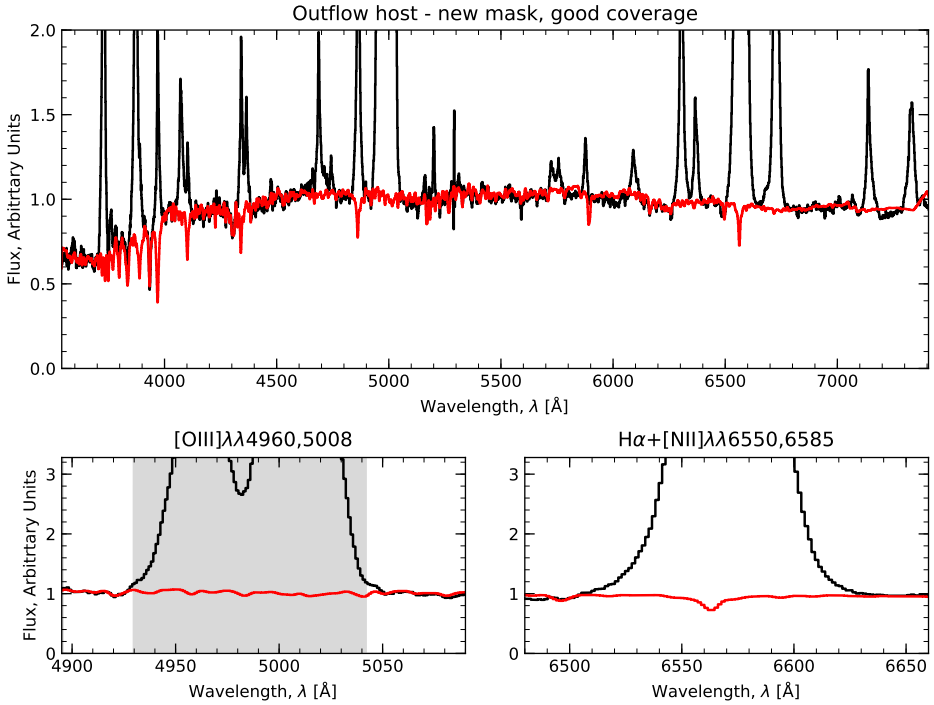
as such the comparison value can be quickly and easily calculated, meeting our computational requirements.

Other spectral decomposition techniques have previously been designed to test for multiple components, such as those employed within `LZIFU` (Ho et al., 2016), that instead considers the Log-Likelihood Ratio Test (LRT) and/or the F-test to decide on the number of components per fit. However, by simply considering the situation in which we recover an accepted fit with $\chi^2_{DoF,2}$=1 (and thus $\chi^2_{DoF,1}$ ~1.11), our test would return p-values in excess of 0.001 for both tests, thus ensuring our fits would be more than 99.9% certain of a second component. Of further note, Protassov et al. (2002) state that, due to non-negative line flux boundary conditions set forth in the emission line fitting procedures, frequentist frameworks such as both the LRT and F-test (and also our χ^2_{DoF} comparison technique) cannot be justified statistically. However, other options, such as a full Bayesian statistical analysis using Monte-Carlo Markov Chain (MCMC) methods, would be too computationally expensive for large datacubes, for instance those of the size of the MaNGA dataset. Furthermore, within Ho et al. (2016), they show that, although maybe not statistically justified, their technique is sufficiently accurate at determining the correct number of components at their required p-values of 0.01. We therefore believe this methodology, which is similar in terms of analysing the best number of fit components is also sufficiently accurate to the more stringent levels we require.

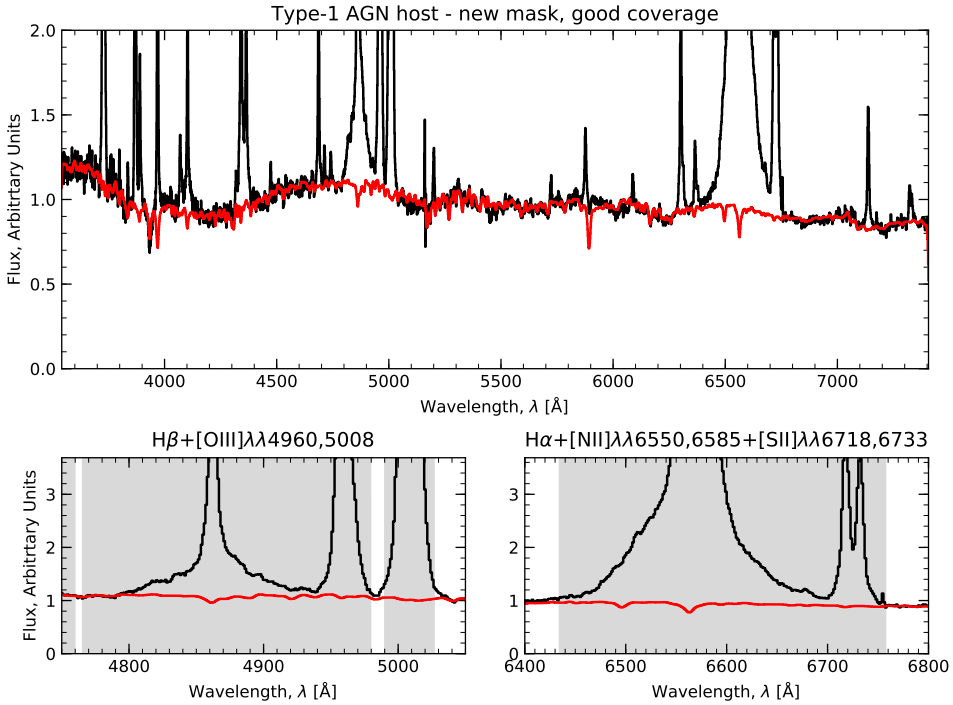
3.6.4 Reassessment in the presence of a broad component

Having recovered the emission lines, SCRAPPy now examines whether the initial mask was sufficiently extensive to cover the wings of the recovered broad component. To assess this, SCRAPPy checks if, for the broad component, the central line velocity plus three times the dispersion is greater than the initial mask width, i.e. by default, $v_b + 3\sigma_b > 1000$ km s^{-1}. If the lines were sufficiently masked (e.g. figure 3.10), the stellar kinematics are assumed robust, and are held fixed from this point onwards. Conversely, if the lines are found to exceed the initial mask, a new mask is constructed, for all emission lines, covering the spectral channels that fall within the afforementioned range. In the case of a type-1 AGN component presence the lines are assumed not to be properly masked, and new masks are reconstructed around all emission lines, with extremely broad masks around the Balmer series, down to a user defined minimum species (by default, the first 4 lines in the series are masked). These masks are used for a new stellar component fit, assessed using the same methodology as previously discussed.

As both this stage and the initial emission line assessment are likely affected by the contamination of the broad emission line wings on the stellar continuum, these last two stages are repeated iteratively, minimally twice. Following this first refit, the lines are re-assessed, the new masks reconstructed and used to recover the emission lines. Subsequent iterations are



(a)



(b)

Figure 3.12: The same two galaxies as in figure 3.11 are shown here, but with good continuum recovery due to the extended masks determined form the initial emission line fits.

only performed if the newly recovered broad component emission line widths are broader than their masks. Upon completion of these iterations, we accept the stellar kinematics as robust and hold them fixed in all subsequent stages, whilst the recovered LOSVDs of all emission line components are provided as initial estimates. Figure 3.12 shows the stellar continuum recovery using these iterative stages for the galaxies shown to have poor initial recovery in figure 3.11. The top panel shows the new continuum recovered for the outflow host galaxy with extremely broad emission line wings, where the emission lines are minimally (yet fully) masked, with recovery around the [OIII] line visually significantly better than that of the initial fit. The lower panel shows the recovery of the type-1 AGN host. In this case, much more of the continuum must be masked due to the presence of the broad Balmer emission lines, especially in the bluer regions of the spectrum where many stellar features are located. However, the fit is again significantly better than the initial fit due to these improved masks.

3.6.5 Simultaneous Fitting of the Stellar Continuum and Emission Lines

The advantages of simultaneously fitting the stellar and emission line components have been explored previously, through a number of different methodologies. This simultaneous fitting approach was first successfully explored by `GANDALF` (Sarzi et al., 2006), wherein the emission lines were modelled as Gaussians and added to the stellar spectra within the code. Sarzi et al. (2006) found that, for emission lines that exist above strong absorption features, the emission line fluxes could be under-estimated if not fitted simultaneously as the absorption features present within the templates could not be modelled due to the removal of the information around/below the masked emission lines. These results were found to most significantly affect weak emission lines, such as Hβ, in galaxies exhibiting lower line emission. As we expect some outflows to exhibit only weak line emission for the broad component in certain line species (e.g. the [NII] and [SII] doublets), a simultaneous fitting approach favours recovery of these lines. These results lead to the improvement of `pPXF` to handle the stellar continuum and emission lines simultaneously, in turn influencing our decision to use `pPXF` as our main fitting engine in SCRAPPy.

The approach taken in SCRAPPy for the simultaneous fitting uses two stages: initially, a three-step iterative process is used to simultaneously fit the continuum and the "major" emission lines, before a final simultaneous fit is performed to recover the final stellar continuum using all of the requested emission lines.

In the initial three-stage iterative process, emission line templates are constructed as Gaussian profiles with an amplitude of unity, centred on the rest wavelength of each "major" emission line and with a width defined by the instrumental resolution. In the case of emission line doublets, a single template is provided for both lines, with the minor line of the doublet modelled

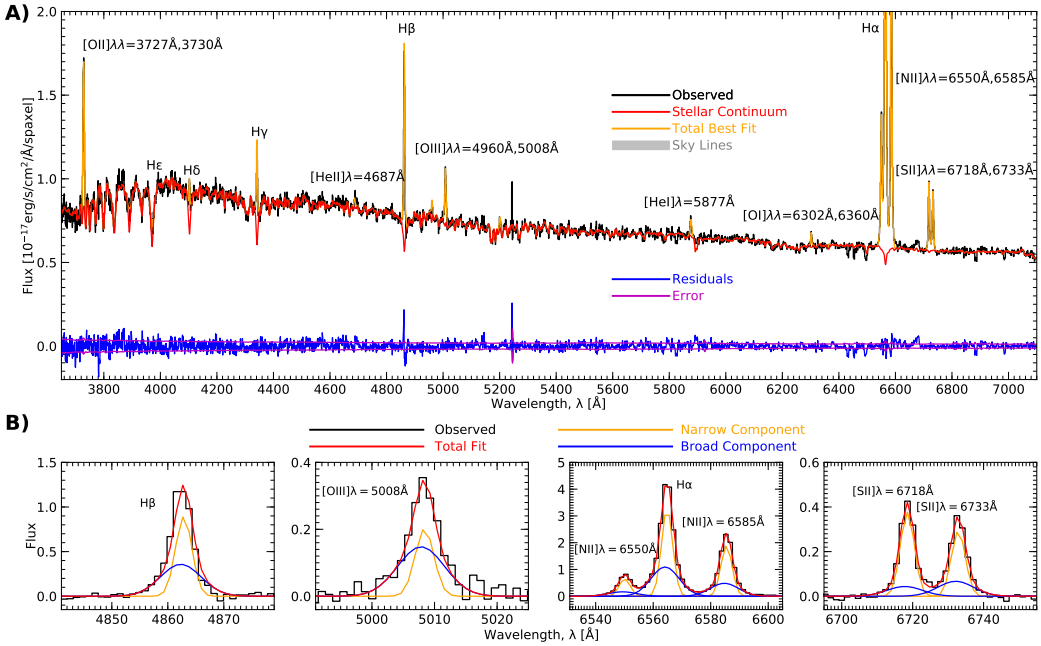


Figure 3.13: Example of simultaneous stellar continuum and emission line fit to the galaxy spectrum shown in figure 3.10. A): Full spectrum decomposition showing the simultaneous stellar continuum (red) and emission line (orange) fit to spectrum. Shown below the fit are the residuals (blue) and the error spectrum (magenta) for the galaxy. The OII 5577 sky line has been masked, shown here by the grey shaded area. B): Subsections of continuum-subtracted spectra, showing some of the brightest nebular emission lines. The decomposition of the narrow (orange) and broad (blue) components can be seen, alongside the total fit of the two components (red).

also as a Gaussian but with an amplitude given by the user provided Einstein co-efficient of the line ratio. If no resolution has been provided, the emission lines are instead constructed at a width of $1/\sqrt{2}$ of the width of the LOSVD recovered by previous stages. `pPXF` is then provided with all of the previous ingredients discussed, with the stellar component kinematics held fixed at the previously recovered values and the initial emission line component kinematics provided as the previously recovered LOSVD (or in the case of no provided instrumental resolution, the recovered velocity but only $1/\sqrt{2}$ of the width). Having fit both the stellar and emission line components simultaneously, a residual spectrum is created by removing the stellar continuum and the emission lines are refit by the emission line fitting module to re-asses their LOSVD. This completes the first iteration of the stage, with the newly recovered emission line LOSVD being passed back in and the analysis being repeated.

Once all three iterations have passed, we assume our emission line fitting has now recovered the emission line LOSVD to sufficient accuracy, and this is used in the second step in the simultaneous fitting: the fitting of the stellar continuum and all emission lines, not simply the

"major" ones. In this stage, the emission lines are constructed at the newly defined central (velocity shifted) wavelength of the line rather than the rest wavelength, with their width set to the recovered velocity dispersion. These templates are also broadened to the instrumental resolution in a similar fashion to the template resolution matching, should the resolution be provided. Whilst the "pixelation" effect previously mentioned will still affect this final template fit, it is significantly lessened by constructing the templates at (or close to) their "true" LOSVD. All kinematics are held fixed at this point, with `pPXF` able to vary only the template weightings. The stellar continuum recovered here is considered final.

3.6.6 Final Emission Line Analysis and Calculations of Properties

Following the simultaneous fit, we perform a final emission line fitting to remove any final impact of the pixelation, and to recover emission line properties needed to produce 2D maps. The stellar continuum (and thus also the stellar absorption features) is (are) assumed at this point to be accurately recovered, thus these fits will have no further impact on the continuum. For each emission line, we know the observed Doppler velocity, compared to the rest wavelength at the heliocentric redshift of the galaxy, and the intrinsic width, σ_{int}. We correct this to the observed total dispersion, σ_{line}, using equation 3.13, calculating the final emission line flux as the integral of the final Gaussian fit:

$$F = A\sigma_{line}\sqrt{2\pi} \tag{3.16}$$

where A is the amplitude of the fit Gaussian. We further calculate the equivalent width and SNR (again calculated as the peak intensity (A) divided by the median error within the FWHM of the line's observed central wavelength).

3.6.7 Stellar Population Synthesis

Feedback from outflow driving sources (both SNe and AGN) can have an impact on the SF of a galaxy. One of the best ways to investigate the role this feedback is currently exerting on a galaxy, its stellar population and the ongoing star formation is through stellar population synthesis. This is achieved by performing a fit using a template library of synthetic populations, such as SSPs, with known characteristic properties that vary from template to template (population to population) such as age and metallicity, and form a regularly spaced grid in this property space. Due to the minimum chi-squared approach used by `pPXF` to converge on a solution, assigned template weights are independent of those assigned to their property-space neighbour, leaving noisy and arbitrary solutions. This degeneracy is not specific to the implementation of `pPXF`, but is a generic mathematical property of the inverse problem. This inverse problem of recovering

the population is ill-posed, but regularisation is a standard approach to solve ill-posed problems (see Cappellari (2017), section 3.5) and is implemented within `pPXF`. This serves to produce smooth maps of the template weight in property space.

To regularise the fit, a certain degree of regularisation must be defined for SCRAPPy to use. The user may define this level in one of three ways: firstly, the user may simply set the regularisation degree in the input files for SCRAPPy; secondly, the user may request SCRAPPy perform an iterative fit to a single spectrum from the galaxy (either the total integrated spectrum or one extracted from the central region) to determine the required level; finally, the iterative fit may be performed on each spectrum, though this method is both incredibly computationally expensive and time inefficient. For the iterative methods, an initial fit is first performed with no applied regularisation, before subsequent fits are performed by increasing the amounts of applied regularisation until the χ^2 of the regularised fit exceeds that of the initial non-regularised fit by a factor of $\sqrt{2N}$, where N is the number of unmasked pixels. This process ensures we end up with the smoothest distribution of weights in property space, whilst remaining statistically consistent with the data.

3.7 Additional implemented features

3.7.1 2D Median Filter Smoothing

As both our emission line fitting module and `pPXF` utilise the Levenberg-Marquardt least-squares method, implemented though the `mpfit` package, fits can be highly biased by initial estimates. To counter this, SCRAPPy offers a final refitting module wherein smoothed maps of the stellar LOSVD, as well as both LOSVD and flux maps for each emission line component, are fed back into the pipeline, and used as initial estimates. This is achieved by spatially smoothing the maps using a 2D median filter. The size and shape of this can be modified by the user, but by default the filter is circular, with a width of 5 pixels. This, coincidentally, equates to the spatial PSF FWHM of both MaNGA and MUSE (though each pixel equates to a different physical size in these datasets).

The use of this module is justified based upon the fact that velocity, velocity dispersion and flux maps typically vary smoothly over the spatial dimensions of a galaxy. This is an effect of the intrinsic properties of galaxies, combined with the finite resolution of the data and smoothing of features caused by the PSF of the telescope used for observations. Hence, initial estimates for these properties can be inferred from the information of neighbouring spaxels.

3.7.2 Implementing Electron Density Ratio Limits

To investigate the electron densities of both the outflowing gas and the gas in the disk of the host galaxy, one can make use of optical line intensity ratios. The primary line doublets used for this technique are the [OII]λ3730/[OII]λ3727 ratio and the [SII]λ6718/[SII]λ6733 ratio (Osterbrock & Ferland, 2006; Pradhan et al., 2006; Proxauf et al., 2014). Both are sensitive to a range of densities generally estimated to be found in both gaseous nebulae and outflows for ionised gas, typically 40 cm^{-3} $\lesssim$ n_e $\lesssim$ 10,000 cm^{-3}. Above and below these limits, the intensity ratio saturates, and densities cannot be accurately recovered.

Of consideration is that the former of these ratios is often difficult to observe, due to the small separation between the two lines. Indeed, for the MaNGA spectra, the [OII]$\lambda\lambda$3727,3730 doublet is blended due to the limited spectral resolution. By default, SCRAPPy will constrain both doublet ratios throughout all fits. However, each doublet will provide an independent measure of the density, as the densities (and thus also relative intensity ratio differences) are not tied between the two doublets. The user may also request other lines are constrained by editing the input file to contain the upper and lower intensity ratios between the lines.

To ensure the intensities of the doublets are constrained to not exhibit physically unsound ratios, we employ two techniques. Firstly, when fitting the emission lines following continuum subtraction, SCRAPPy fits the lines typically defined as the numerator (e.g. [SII]λ6718) in the intensity ratio calculations with no constraint to their amplitudes. Using these amplitudes, SCRAPPy defines the amplitudes of the numerator lines (e.g. [SII]λ6733) via multiplication with a factor constrained between the density limits. Secondly, when fitting the stellar continuum and emission lines simultaneously using `pPXF` we provide two templates for each doublet, sampled at the upper and lower density limits. `pPXF` attempts to fit the gas lines efficiently and accurately as a linear least-squares problem, whereas implementing limits would make the solution a slower, less reliable general non-linear one. By using two templates, the final fit results as a linear combination of the templates, and the emission line peaks will be constrained by the templates.

3.7.3 D4000 and D$_n$4000 Indices Calculations

Spectral indices can provide us with information about the properties of the underlying stellar population. One of the key spectral features that has been studied in depth are the ubiquitous D4000 and D$_n$4000 indices that trace the strength of the 4000Å break, which arises due to a combination of the Balmer break and absorption by various metal lines, depending on the age and metallicity of the stellar population. Both D4000 and D$_n$4000 are defined as ratios between two standard bandpass integrals, taken redward and blueward of 4000Å. The D4000 feature is

	Band Passes (Å)	
Index	Blue	Red
D4000	3750-3950	4050-4250
D_n4000	3850-3950	4000-4100

Table 3.3: Bandpass limits for both D4000 and D_n4000

defined as

$$D4000 = \frac{\int_{4050(1+z)}^{4250(1+z)} f_\nu d\lambda}{\int_{3750(1+z)}^{3950(1+z)} f_\nu d\lambda} \tag{3.17}$$

with D_n4000 being similarly defined but with narrower bandpass limits to mitigate the effects of reddening. We note here that, as SCRAPPy shifts spectra to their rest frame prior to fitting, the (1+z) factor in both integrals will disappear. The bandpass wavelength limits for both indices are shown in table 3.3. However, in this definition, the flux being integrated over is given per unit frequency, f_ν, whereas most spectroscopic data fluxes are given per unit wavelength, f_λ. Further, the spectra are sampled at discrete pixels, thus we must redefine these bandpass integrals as discrete sums. Hence, the each integral in equation 3.17 is redefined as the discrete sum

$$\int_{\lambda_l}^{\lambda_u} f_\nu d\lambda = \frac{1}{c} \sum_i \delta_i \lambda_i^2 f_{c,i} d\lambda_i \tag{3.18}$$

where λ_u and λ_l are the upper and lower wavelengths of the band passes respectively, c is the speed of light, $f_{c,i}$ is the emission subtracted stellar continuum in pixel i, centred at wavelength λ_i with width $d\lambda_i$, and δ_i represents the fractional amount the pixel is within the bandpass, defined as

$$\delta_i = \begin{cases} 1 & \text{both pixel edges fully within bandpass} \\ 0 < \delta_i < 1 & \text{only one pixel edge within bandpass} \\ 0 & \text{otherwise} \end{cases} \tag{3.19}$$

In this way, both the D4000 and D_n4000 are calculated from the ratio of two discrete sums. We calculate both indices as each has pros and cons for their uses, and calculate each for both the observed spectrum, with the best fitting emission lines removed, as well as for the best fitting stellar continuum, constructed from the templates by `pPXF`.

3.8 Data outputs

Previously touched upon, SCRAPPy produces outputs of all initial assessments of the data (SNR, binning, etc.), and of all spectral fits and calculated properties. The final step of

SCRAPPy is to consolidate all of the spectral fitting data into two datacubes containing the 2D property maps (e.g. v, σ) and all components of the spectral fit (e.g. the raw spectrum, the `pPXF` bestfit, polynomials). All of the data has been stored in such a way that the fits at any point in the pipeline can be reconstructed, such that any encountered errors can be targeted and rectified.

This consolidation of data is an important step for any pipeline that aims to reduce IFU data. Whilst each of the spectral fittings is available and saved, and the user could reconstruct cubes themselves from the relevant 2D spatial information, ease of use for the user is paramount and established with this consolidation. Of these constructed consolidated datacubes, the most relevant to almost any user will be the 2D maps of stellar and emission line properties. These 2D maps are stored as multi-extension fits files, with extensions containing either single maps, for properties such as the stellar LOSVD moments, or a number of maps stored in channels along the 3^{rd} axis, with the extension header detailing what is contained in each channel. The stored flux values are not extinction corrected, and are given as flux per spaxel.

The final spectral fits datacube contains the raw binned/spaxel galactic data, alongside all generated masks, `pPXF` outputs (e.g. bestfit, polynomials) from the final simultaneous fit and constructed emission line profiles from the final emission line fit, all within separate 3D extensions. For binned spectra, the fit is saved in each constituent bin spaxel location. Whilst most users will not need to explore these files, they are useful to examine residuals around the emission lines for extreme objects.

3.9 Conclusions

Here we have presented an overview of a new pipeline, SCRAPPy, for the express purpose of reducing 3D (IFU) spectral cubes of galaxies where multiple emission line components are needed, such as those galaxies hosting outflows, type-1 AGN or merging systems. Testing the accuracy of this pipeline is a difficult task, as there are few other pipelines capable of reducing this type of data. However, this pipeline has assisted in two studies of outflows on data from the MaNGA survey, and an early version of this pipeline was used in the first detection of positive feedback, using X-shooter data, proving both its flexibility and its ability to sufficiently recover the data to the degree of accuracy needed for study.

Whilst some effort has gone into determining properties of the underlying stellar population through stellar population synthesis, and calculations of both the D4000 and D_n4000 indices are performed, the main design philosophy of SCRAPPy has been to focus on multiple component detection and emission line recovery. Later versions of SCRAPPy will focus more upon the recovery of the stellar information to investigate the effect feedback from outflows may have on

the stellar population and its prior evolution.

The emission line fitting module has also only been fully tested for fits to galaxies with two emission line component and, though it has proven robust enough to do so, has not yet been used to analyse type-1 AGN. Again, later versions of SCRAPPy will focus upon extending the ability of the user to allow more than two emission line components, especially for spectra with high spectroscopic resolutions (e.g X-shooter), for which it will be easier to disentangle a true additional emission line component from background noise or fitting artefacts. We also wish to add the ability at a later date to include complex spectra observed of AGN, containing features such as dense metal emission line forests and bright auroral lines with complex profiles.

Chapter 4

Star formation inside a galactic outflow

This chapter is adapted from both 'Star formation inside a galactic outflow', R. Maiolino et al. 2017, for which the author (R. Gallagher) provided significant contributions towards data reduction, data analysis and editorial support.

As discussed in the introduction, recent observations have revealed massive galactic molecular outflows (Feruglio et al. 2010; Sturm et al. 2011; Cicone et al. 2014) that may have the physical conditions, such as high gas densities (Aalto et al. 2012; Aalto et al. 2015; Sakamoto et al. 2009), required to form stars. Several recent models predict that such massive outflows may ignite SF within the outflow itself (Ishibashi & Fabian 2012; Zubovas et al. 2013; Silk 2013; Zubovas & King 2014; Ishibashi et al. 2013, Decataldo et al. 2019) This SF mode, in which stars form with high radial velocities, could contribute to the evolution of the spheroidal component of galaxies, and could have several other important implications (discussed in chapter 5). However, until recently, there was no observational evidence of this potential new mode of SF. In this chapter, I report the first discovery of SF inside a galactic outflow in the galaxy IRAS F23128−5919, obtained through observations performed with MUSE and X-shooter at the VLT.

4.1 Target and archival MUSE data

IRAS F23128−5919 is a merging system (figure 4.1(a)), in which the southern nucleus hosts an obscured active galactic nucleus (AGN), detected in the X-rays (Brightman & Nandra, 2011).

Past observations had already revealed a prominent outflow developing from the southern nucleus (Piqueras López et al. 2012; Bellocchi et al 2013; Leslie et al. 2014; Arribas et al. 2014; Cazzoli et al. 2016), driven by the nuclear starburst, or by the AGN, or both. We analysed archival Very Large Telescope (VLT) spectroscopic observations, obtained with the MUSE instrument, of the optical nebular lines to better characterise the outflow. MUSE is a large IFU spectrograph covering the spectral range 4750Å<λ<9300Åwith a spectral resolution varying from R=2000 to R=4000 across the wavelength range. The data of IRAS F23128−5919 was obtained with an on-source integration time of 52 min. The seeing during the observations was only about 2.5″.

The stellar continuum was fitted and subtracted at each pixel by using the `pPXF` routine (Cappellari 2017) by using the MILES templates (Sanchez-Blazquez et al. 2006). The nebular lines Hα, [NII]$\lambda\lambda$6548,6584, [SII]$\lambda\lambda$6717,6731 [OII]λ6300 and [OIII]$\lambda\lambda$4959,5007 were simultaneously fitted with two Gaussian components: a narrow one with FWHM < 300 km s^{-1} and a broad one with FWHM > 300 km s^{-1}. At each spatial position the velocity and velocity dispersion of each of the two components was fixed for all nebular lines. The intensities of the [NII] doublet and of the [OIII] doublet were tied to have the ratio given by their Einstein coefficients. The resulting maps of the Hα flux, velocity and velocity dispersion of the two components in figure 4.1(a) are obtained by masking out regions for which the SNR of the component is lower than five. We note that in previous work (Bellocchi et al. 2013), which analysed data of the same source, obtained with the VIMOS instrument at the VLT, the most Eastern nebular emission was associated with the 'narrow' component but fitted with very large velocity dispersion ($\sigma_v \approx$ 300 km s^{-1}) and was very blueshifted (v < -300 km s^{-1}). This mismatch was probably due to the lower SNR of the VIMOS data in those external regions, which made the line decomposition much more difficult. In our analysis, using MUSE data, we have corrected this (not completely appropriate) association: such high values of the velocity dispersion and strong blueshift should actually be associated with the 'broad' blueshifted component which traces the outflow.

The narrow component (bottom row in figure 4.1(a)) is probably tracing the bulk of the dynamics of the two merging disks. The outflow traced by the broad blueshifted component (top row in figure 4.1(a)) of the nebular lines extends towards the East of the southern nucleus for about 7−9 kiloparsecs (8″-10″), beyond the optical galactic disk. A receding counter-outflow is also observed in the opposite direction, that is, towards the West (although it is weaker, owing to extinction by the galactic disk).

4.2 New X-shooter data

We have observed the central and eastern outflows with the X-shooter spectrograph at the VLT. X-shooter is a long-slit medium-high resolution echelle spectrograph with a single slit that simultaneously observes in the wavelength range 3500Å<λ<2.5μm. We selected the 11″-long slit with a width of 1.3″ in the ultraviolet B (UVB) arm (wavelength spectral range 300−560 nm), 1.2″ in the visible (VIS) arm (spectral range 550−1,020 nm) and 1.2″ in the near-infrared (NIR) arm (spectral range 1,020−2,480 nm). The resulting spectral resolutions are R = 4,000 in the UVB, R = 6,700 in the VIS and R = 3,890 in the NIR. The seeing during the observations was about 0.9″ on average. The slit was positioned with a position angle of 70° with the southern nucleus centred at 3.2″ from the centre of the slit (figure 4.1(a)); this orientation and centring enables the slit to sample both the nuclear and the Eastern part of the outflow. Observations were executed by nodding the slit by a few arcsec (to optimise the removal of detector artefacts) and interleaved with blank sky exposures, obtained at about 1 arcmin from the galaxy, for sampling the background. The total on-source integration time was 1.7 h.

Data reduction and calibration began by following the standard pipeline steps (Modigliani et al. 2010). Great care was taken to correct the atmospheric absorption features, not only in the NIR, but also in the VIS, by using the telluric standard and by smoothing its spectrum to the resolution of the science observations (the spectral resolution of the telluric standard is much higher, by about 30%, than the science observation because the seeing was smaller than the slit). We extracted six spectra along the slit, with variable apertures (from 0.5″ to 2.5″), to find a trade-off between spatial information and the signal-to-noise ratio.

The fitting was performed using a prototype version of SCRAPPy, using the MILES synthetic library (Sanchez-Blazquez et al. 2006). We employ the full library in the fittings. However, we note that, unfortunately, templates covering the required spectral range at the high resolution of our spectra are available only for ages down to 30 Myr. This constraint implies that we cannot disentangle populations younger than 30 Myr. More specifically, all spectral fits require a population of 30 Myr (jointly with contribution from older stellar populations), but this must be regarded as an upper limit of the youngest stellar population because of the library limits. The standard SCRAPPy procedures were used, however the He I emission and absorption line were fitted (post fit) separately, as this feature may trace a different stellar population with a different kinematics. The Ca II triplet was also fit separately to the rest of the fit, and was restricted to the bluest of the lines (λ=8948Å) as the other two lines are heavily affected by sky emission lines and telluric absorption.

In this fitting, we considered having SCRAPPy allow the different stellar populations in the stellar fit to have different velocities. In particular, we considered in principle allowing the old

stellar population to have a different velocity relative to the young stellar population. However, this was not really possible in practice because: (1) this additional degree of freedom would introduce a lot of degeneracy when also attempting to simultaneously deblend the nebular line emission from the stellar features, and (2) the young stellar population dominates the observed stellar light across the entire spectrum, hence the velocity traced by the spectral features associated with the old stellar population is extremely difficult to recover. Note that the signal-to-noise ratio in the spectra allows us to trace the stellar kinematics through the He I line only out to 1.2 kpc, and through the Ca II triplet only out to 2.3 kpc. At 5 kpc the stellar kinematics can be recovered only through the Balmer absorption lines. At 6.6 kpc (the outermost point at which the spectrum is extracted) the very low signal-to-noise ratio on the continuum does not allow us to detect stellar signatures. It is interesting to note that the spectrum extracted at 2.3 kpc (sampling from 2″ to 3.6″ from the nucleus) includes a clear knot observed in the HST image. This could be a star cluster that has formed within the outflow. Indeed, in the position-velocity diagram (figure 4.4(b)) it is located on the ballistic track of particles formed in the inner part of the outflow.

Regarding how this prototype SCRAPPy handled the fitting of the nebular lines, we allowed up to 3 emission line components. In the optical spectra, the higher spectral resolution and higher SNR of the X-shooter data, relative to the MUSE data, reveals that the broad blueshifted component tracing the outflow cannot be reproduced with a single Gaussian. In the central aperture the broad component requires at least two blueshifted broad Gaussians (blue lines in figure 4.1(b)) to be properly fitted (along with a narrow component accounting for gas in the host galaxy). In some of the Eastern apertures, dominated by the outflowing gas, the broad blueshifted component requires three Gaussian components to be properly fitted. In the latter apertures a very weak narrow component (even more clearly separated from the broad component), associated with the outskirts of the host galaxy in the background is also always detected. In the central aperture, absorption by a diffuse interstellar band at λ = 6,282 Åaffects the blue shoulder of the bluest and broadest component of the [OI]λ6,300 line, so the spectral region in the range 6125 Å < λ_{rest} < 6,291 Å was not included in the fit. This results in a slightly higher uncertainty of the fit parameters for this component of the [OI] line

In the NIR data of the nuclear region (which is the only region where we can extract the spectrum with a signal-to-noise ratio high enough for a proper spectral decomposition), the profile of the nebular lines changes substantially relative to the optical lines. This is primarily due to the effect of differential dust extinction, as a consequence of which some components are much more absorbed in the VIS than in the NIR. Therefore, we fitted the NIR lines separately with a different set of components, not tied to the parameters of the optical components, although the general properties are similar.

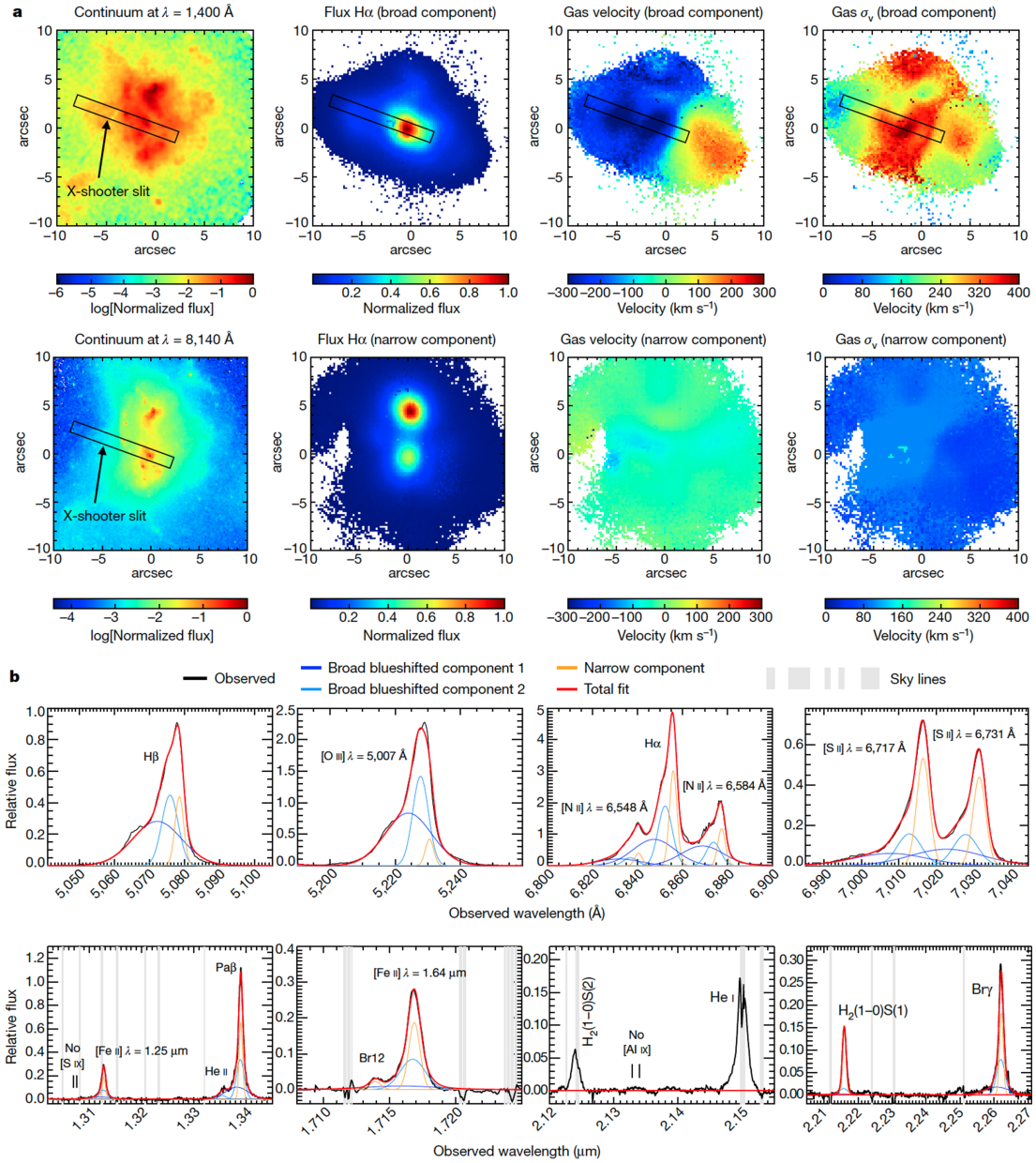


Figure 4.1: Spectral decomposition of the outflow in IRAS F23128-5919. (a) HST continuum images, surface brightness (flux per surface area), velocity and velocity dispersion σ_v maps of the broad and narrow components of the Hα line, inferred from the MUSE data. The location and orientation of the X-shooter slit is shown (boxes). (b) Subsections of continuum-subtracted X-shooter spectra, extracted from the central region, around some of the relevant emission lines, showing the decomposition between narrow and broad components, as well as the non-detection of coronal lines.

The slit location (box) is shown in figure 4.1(a). An example of a spectrum extracted from the central region is shown in figure 4.1(b), zoomed around some of the relevant emission lines.

4.3 Data analysis

4.3.1 BPT diagrams

For the optical nebular lines, we corrected for dust reddening, using a Milky Way extinction curve (Cardelli et al. 1989), by using the Hβ/Hγ ratio to infer the reddening. Hα/Hβ cannot be used, as these two lines were observed with different apertures.

Figure 4.2(a) shows the location of the various emission-line components tracing the outflow on the so-called 'BPT diagnostic diagrams', which are widely used to obtain an approximate discrimination between different excitation mechanisms (Kewley et al. 2006, and references therein). Most components describing the outflow (broad blueshifted components) are located in the region of the diagram populated by star-forming galaxies and HII regions, with only a few exceptions mostly restricted to the diagram involving the [NII] line; however, this diagnostic has potential problems associated with the nitrogen abundance (Pérez-Montero et al. 2009; Shapley et al. 2015) and, in this specific target, the [NII] line is also affected by telluric absorption, whose correction introduces additional uncertainties.

The finding that nebular components tracing the outflow are located in the star-forming locus of the BPT diagrams provides a first indication that SF is probably occurring in the outflow. However, the BPT diagrams do not necessarily provide an unambiguous classification. For instance, there are excitation models that predict that some shocks could produce line ratios in the 'star-forming' region of the BPT diagrams (Allen et al. 2008).

4.3.2 NIR diagnostics

Fortunately, the X-shooter wide spectral range provides a wealth of additional diagnostics of the excitation mechanism. The NIR iron lines are extremely useful to identify shock excitation. Indeed, shocks are known to destroy dust grains, hence releasing into the interstellar medium large amounts of iron, which is otherwise locked into grains. This results in much higher emission of iron NIR transitions, relative to the hydrogen recombination lines, than is typically observed in star-forming regions. Shock heating of the interstellar medium also results in excitation of the NIR vibrational transition of molecular hydrogen. As such, Shock/AGN excitation can be cleanly distinguished from excitation by young stars in the [FeII]λ = 1.64 μm/Brγ λ = 2.16 μm versus the H_2(1−0) S(1) λ = 2.12 μm/Brγ diagram (Colina et al. 2015), as shown in figure 4.2(b), where the measurements of IRAS F23128−5919 (symbols) are

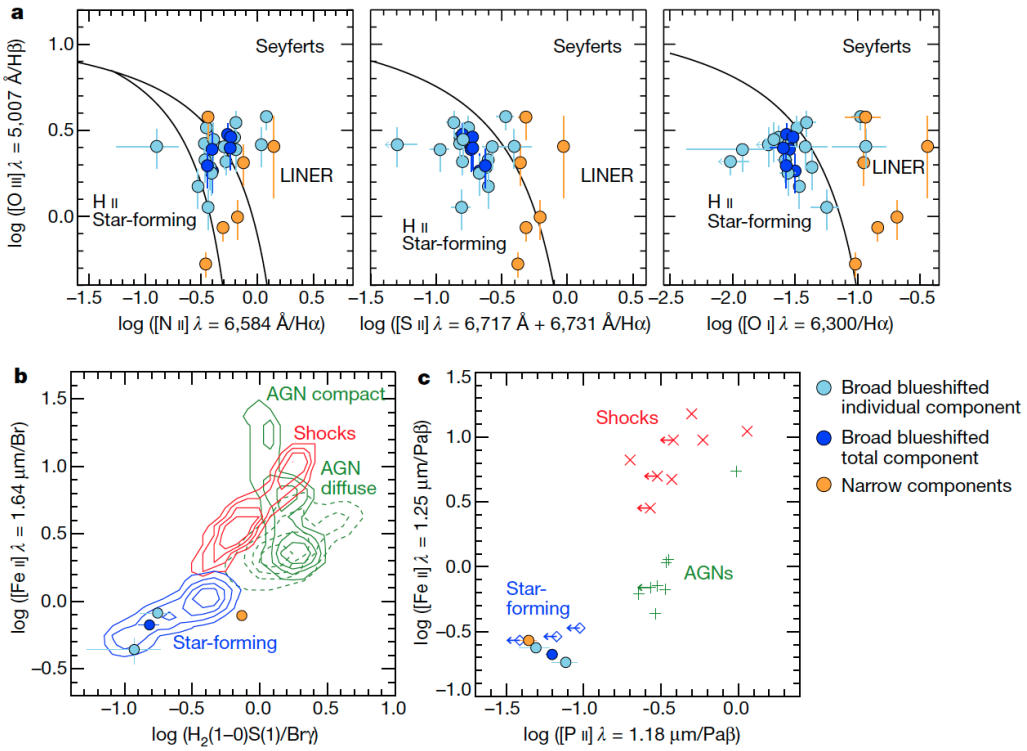


Figure 4.2: Diagnostic diagrams. Distributions of the individual broad Gaussians (light-blue symbols) and of the resulting total broad components (dark-blue symbols) tracing the outflow, as well as of the narrow components (orange symbols) tracing gas in the host galaxy, extracted from different apertures along the X-shooter slit. a, BPT diagnostic diagrams involving optical nebular lines, showing the curves defining the HII, Seyfert and low-ionization nuclear emitting regions (LINERs) loci.

compared with the locations of known HII, AGN and shocks, shown in contour. Blue symbols indicate the line ratios of the components tracing the central part of the outflow in IRAS F23128−5919, indicating that the gas excitation in the outflow is consistent with SF and not consistent with other excitation mechanisms, such as shocks or AGNs.

Furthermore, the low ionisation line [PII] $\lambda = 1.18$ μm is an excellent diagnostic of partially ionised, transition regions produced by X-rays, produced either by radiative shocks or by AGNs (Oliva et al. 2001). Hence, The [FeII] $\lambda = 1.25$ μm/Paβ versus [PII] $\lambda = 1.18$ μm/Paβ diagram is also a good discriminator of the excitation mechanism (Oliva et al. 2001), as illustrated in figure 4.2(c), where the measurements of IRAS F23128−5919 (symbols) are compared with the distribution of known HII, AGN and shocks. The outflow in IRAS F23128−5919 is inconsistent either with shock excitation or with AGN photoionisation. The outflow line ratios are consistent with those observed in star-forming galaxies.

As illustrated in figure 4.1(b), the spectrum of IRAS F23128−5919 does not show evidence

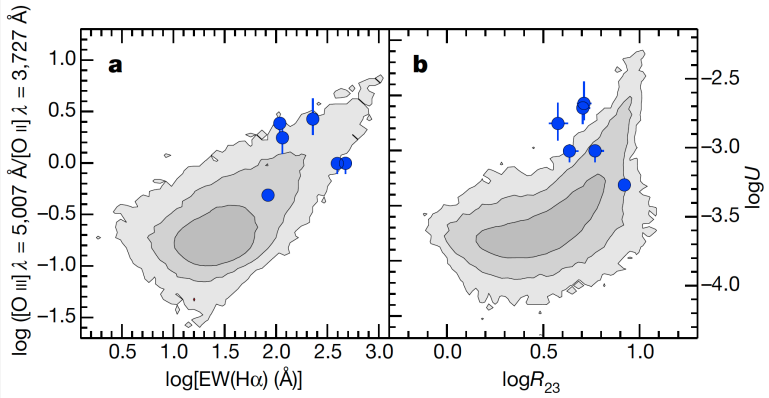


Figure 4.3: Ionisation parameter. The [OIII]/[OII] line ratio (left-hand y-axis) is sensitive to the adimensional ionisation parameter U, as given on the right-hand y-axis (although it also has a secondary dependence on metallicity). (a) [OIII]/[OII] as a function of equivalent width EW(Hα), which is used as a proxy of the age of the stellar population. Contours indicate the distribution of star-forming galaxies from the Sloan Digital Sky Survey survey (68%, 95% and 99.7% of the population), illustrating that younger systems have higher ionization potential. Symbols indicate the location of the gas in the Eastern outflow (we cannot differentiate between individual broad components since we cannot differentiate the continuum associated with each component), indicating that it is consistent with *in situ* photoionization by young stars. (b) The [OIII]/[OII] line ratio has a function of R_{23} = ([OIII]$\lambda\lambda$4960,5008 + [OII]λ3727)/Hβ, which is sensitive to the metallicity, with a secondary dependence on ionization parameter. The diagram further supports the idea that the gas excitation in the outflow does not differ substantially from normal star-forming regions and, if anything, the ionization parameter is even higher. Error bars show the 1σ uncertainties.

for coronal lines, which are generally associated with powerful AGNs. This further supports the absence of a substantial AGN contribution to the gas excitation. Therefore, neither shocks nor the AGN can account for the excitation of the gas in the outflow, whereas all diagnostics are consistent with excitation by SF.

4.3.3 The ionisation parameter

The presence of young stars is clearly revealed in the ultraviolet Hubble Space Telescope (HST) images (figure 4.1(a)). However, from imaging alone it is not possible to disentangle putative young stars in the outflow from young stars in the galaxy disks, whose ultraviolet radiation field can potentially ionise the gas in the outflow (externally) and produce the line ratios observed in the outflow. However, we should be able to discount an external source of stellar ionisation by exploiting the ionisation parameter, defined as ratio between ionising photons flux ($Q_{ion}/4\pi r^2$) and gas electron density (n_e), and normalised through the speed of light,

$$U = \frac{Q_{ion}}{4\pi r^2\, c\, n_e} \tag{4.1}$$

The gas density in the outflow is similar or even higher than in the disk, as inferred from the [SII] doublet (discussed in chapter 6). Therefore, in the scenario of external photoionisation, one would expect that the ionising flux should be much lower than in the case of *in-situ* photoionisation, therefore resulting in an ionisation parameter orders of magnitude lower than observed in standard star-forming regions.

The ionisation parameter can be traced through the [OIII]λ5008/[OII]λ3727 line ratio (Díaz et al. 2000). This ratio has however also a secondary dependence on the gas metallicity. The latter can however be monitored through the R_{23} parameter, defined as the ratio ([OIII]$\lambda\lambda$4960,5008 + [OII]λ3727)/Hβ, which is primarily sensitive to the oxygen abundance, and with a secondary dependence on the ionisation parameter. Therefore, diagrams with both these two quantities are used to help disentangling these properties (e.g. Nagao, Maiolino & Marconi 2006).

As illustrated in figure 4.3, the ionisation parameter of the gas in the outflow is not lower than observed in normal star-forming regions, and is in fact slightly larger, as illustrated in figure 4.3. This result strongly argues in favour of *in situ* stellar photoionisation, by young stars within the outflow.

4.3.4 Stellar Population Kinematics

Even better evidence for stars formed in the outflow is the direct detection of a young stellar population with the kinematical fingerprints of formation inside the outflow. The stellar continuum is detected in the X-shooter spectra out to about 5 kpc from the nucleus, and its shape and stellar features do indeed indicate the presence of a young stellar population younger than a few tens of millions of years (Myr). Determining the kinematics of such a stellar population is difficult, since most stellar features (especially the ones associated with young stars) are heavily contaminated by the strong nebular emission lines. However, we have recovered the kinematics of the young stellar populations through our spectral fitting of the optical spectrum (figure 4.4(a)) which is dominated by the Balmer lines (tracing young hot O-B-type stars, but also contaminated by older A-type stars) and through the Ca II triplet (CaT) at wavelength $\lambda \approx$ 8,500 Å (which, in the case of recent SF, is dominated by young red supergiants and young asymptotic giant branch stars, although the latter are also contributed by older stellar populations). Moreover, we have detected (only in the regions of highest signal-to-noise ratio) the weak absorption feature of HeI λ = 4,922 Å (figure 4.4(a)) which is an unambiguous tracer of B-type stars and of stellar populations with age (González Delgado et al. 2005) of about 10 Myr.

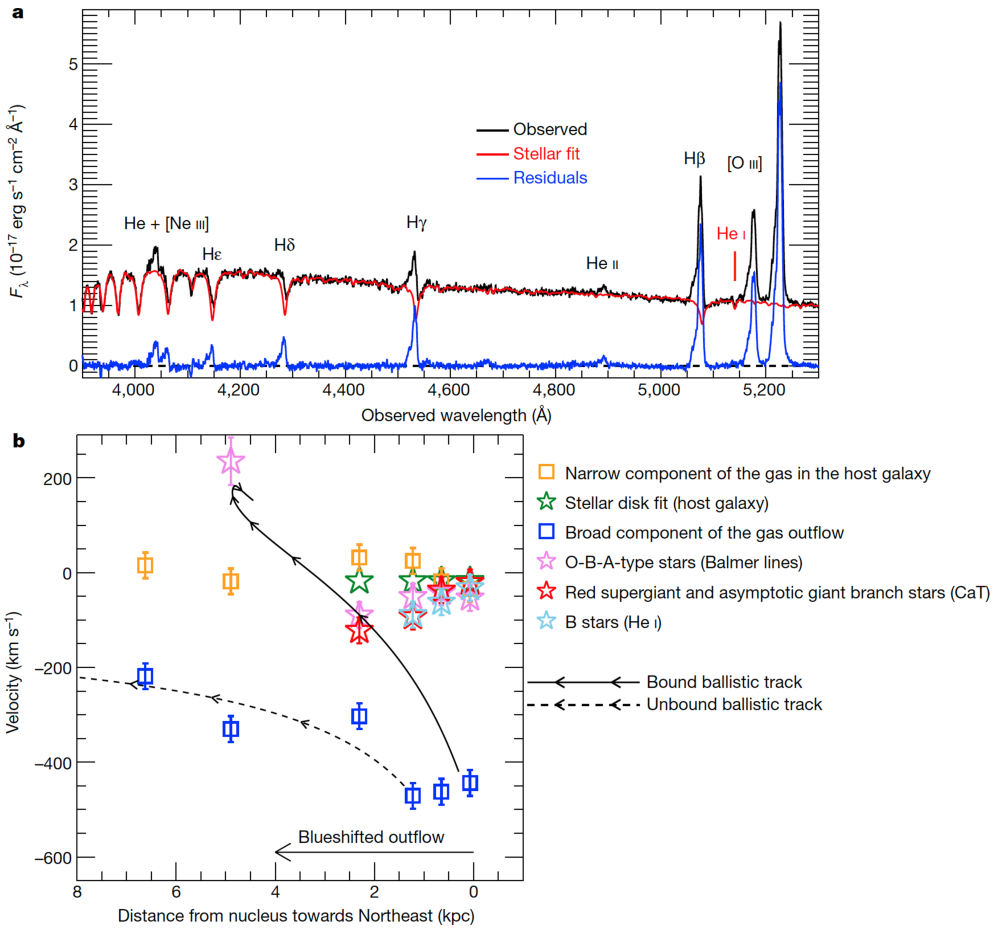


Figure 4.4: Kinematics of the young stars compared with the gas kinematics. (a) Stellar fit of the spectrum extracted from the outflow region at 1 kpc from the nucleus. F_λ is the intensity of the line or flux per unit wavelength. (b) Velocity of the young stellar population along the outflow, traced by different features compared with the kinematics of the broad nebular component tracing the gaseous outflow (centroid of the broad Gaussians fitting the broad component; dark-blue symbols), and the kinematics of the host galaxy disk traced by the narrow nebular component (orange symbols) and by a fit of the disk stellar kinematics in the MUSE data (green symbols). Black curves show the position-velocity tracks of stars formed in the outflow, according to the model described in the Methods. Error bars show the 1σ uncertainties.

Figure 4.4(b) shows the kinematics of the various components as a function of galactocentric distance along the X-shooter slit. Except for the central kiloparsec, where stellar features are probably dominated by the central starburst in the host galaxy, all stellar features in the outflow region, at a galactocentric radius of about 1–3 kpc, show a blueshift relative to the galactic disk, indicating that they are indeed associated with the outflow. The velocity of the stellar features (reaching a maximum blueshift of about 100 km s^{-1}) is much lower than the velocity observed for the nebular lines in the outflow (about 250–450 km s^{-1}); however, this difference

is expected in the scenario in which stars form in the outflow.

Indeed, although the gas in the outflow is probably maintained at high velocity (at least in the central 1.5 kpc) by the continuous action of radiation pressure or ram pressure by expanding hot gas, as soon as stars form in the outflow they react only to gravity, and are rapidly decelerated by the galaxy gravitation field (that is, they move ballistically). A simple dynamical model easily describes such an effect. The gravitational potential within the central few kpc of the galaxy is supposed to follow the axisymmetric Miyamoto-Nagai disk and bulge in the form:

$$\phi(r,z) = -\frac{GM}{\sqrt{r^2 + \left(a + \sqrt{z^2 + b^2}\right)^2}} \tag{4.2}$$

where M is the total dynamical mass of the system, G is the gravitational constant, r is the radial cylindrical coordinate on the plane of the disk, z is the coordinate perpendicular to the disk plane, and a and b are constants (Miyamoto & Nagai, 1975). The curves shown in figure 4.4(b) assume $M = 1.5 \times 10^{11} M_{\odot}$, $a = 1.7$ kpc and $b/a = 0.1$. The dynamical mass may appear somewhat large, and in particular larger than given in previous works (Bellochi et al. 2013); however, one should take into account that (1) in the Miyamoto-Nagai potential the quantity M is the total mass (integrated to infinity) associated with the potential and (2) velocity projection effects make it difficult to infer a solid dynamical mass of the system, especially for nearly face-on systems (in particular the inferred dynamical mass is generally a lower limit, modulo the cosine of the inclination angle). For the sake of simplicity and for the sake of reducing the number of free parameters we have assumed that the galactic disk is oriented perpendicular to the line of sight (this is a reasonable assumption given that the rotation curve of the disk is very weak). In the model the velocity of the outflowing gas makes an angle of about 45° with the line of sight.

In figure 4.4(b) the solid curve shows the position-velocity track of stars formed at the base of the outflow (specifically at 500 parsecs from the nucleus, corresponding to a projected distance of about 350 pc), where most of the SF in the outflow is occurring, as inferred from the broad Hα flux map. Within this model such stars are rapidly decelerated and their velocities even change sign (that is, they fall back towards the disk and towards the bulge), reproducing the positive velocity of young stars observed at 5 kpc from the nucleus. Within this model, stars formed in the outflow at projected distances larger than 1 kpc are only mildly decelerated and become gravitationally unbound (see the dashed curve in figure 4.4(b).

4.3.5 Star formation rates

By using the total Hα emission of the gas in the Eastern (approaching) outflow, and corrected for extinction through the Balmer decrement, we infer a total star-formation rate in the Eastern outflow of about 15 $M_\odot$ yr^{-1}, assuming a 'Chabrier' initial mass function (Kennicutt & Evans 2012). Performing a similar analysis in the Western (receding) outflow is prevented by the large dust extinction due to the galactic disk. However, if the Western outflow is also characterised by similar SF, then the total SFR occurring in the outflow is about 30 $M_\odot$ yr^{-1}, that is, a significant fraction (about 25%) of the total star-formation rate of the merging system (Rodríguez-Zaurín et al 2011), about 115 $M_\odot$ yr^{-1}.

4.4 Conclusions

A number of theoretical models have predicted that SF should occur within massive galactic outflows, with the prediction that these stars formed on ballistic tracks could have far reaching implications in the kinematics properties of their host galaxy. Presented here is a wealth of observational evidence which, combined, supports the scenario in which stars have formed within the outflow of this galaxy.

The individual pieces of evidence are as follows:

(i) The ionised gas within the outflow falls within the SF locus on both the [SII]- and [OI]-BPT diagrams, and in either the SF or Composite loci on the [NII]-BPT diagram. This suggests the major source of ionisation for the gas is UV radiation emitted by young, hot stars.

(ii) The NIR diagnostics of the gas suggest that shock/AGN excitation cannot explain the observed ratios in the [FeII]λ = 1.64 μm/Brγ λ = 2.16 μm versus the $H_2(1-0)$ S(1) λ = 2.12 μm/Brγ and the [FeII] λ = 1.25 μm/Paβ versus [PII] λ = 1.18 μm/Paβ diagrams, instead suggesting excitation from young stars is responsible for the observed ratios.

(iii) UV HST images of the galaxy clearly show the presence of young stars. Their presence in the outflow supported by the ionisation parameter of the outflowing gas, found to be consistent with that of HII SF galaxies from the SDSS-DR7 dataset, suggesting *in-situ* ionisation from young stars.

(iv) Tracing the kinematics of different stellar populations, we can see that the young stellar population have blueshifted kinematics with respect to the stars within the disk.

Having confirmed the presence of SF within the outflow, we estimate that the total amount of SF in the outflow. We calculate that for the Eastern outflow there is 15 $M_\odot$ yr^{-1} of SF. The Western outflow is obscured by dust, so we cannot calculate the SFR in this outflow, but expect

a similar amount. Thus, the total SFR in the outflow is 30 $M_{\odot}$ yr^{-1}, approximately 25% of the total SFR of the galaxy.

Chapter 5

Widespread SF inside galactic outflows

This chapter is adapted from both 'Widespread SF inside galactic outflows', R. Gallagher et al. 2019

Within chapter 4 we have presented the first evidence of SF *inside* a galactic outflow for a specific galaxy studied in detail. Whilst this has been a remarkable result, it has not been clear whether this is an isolated, rare case, or SF is more common in galactic outflows, but it has been difficult to identify.

Indeed, outflows are often AGN-driven, which (as we shall discuss in Section 5.5) tend to dominate the diagnostics, especially in integrated or nuclear spectra, and the presence of shocks may also hinder the detection of SF.

We have investigated the occurrence of this phenomenon in a systemic and unbiased way by exploiting the integral field spectra of the 2800 galaxies of the MaNGA-SDSS4 DR2 survey. These offer high quality, spatially resolved spectra spanning a broad wavelength range, hence enabling us to probe multiple diagnostics. We identify a subsample of 37 galaxies that show evidence for clear outflows and which can be traced through all primary nebular lines required for identifying the sources of gas ionisation/excitation through diagnostic diagrams. We have therefore exploited these data to map the diagnostics across each outflow, hence revealing the presence of SF inside outflows. As we will discuss in detail, we have found that about one third of outflows are characterised by prominent SF inside them, and about half of the outflows show evidence for at least some SF, hence revealing that SF is common to most galactic outflows, with major implications especially for high-redshift galaxies, where outflows are much more

prominent.

5.1 Data analysis

5.1.1 Data reduction

We considered all IFU spectra of the 2800 galaxies in the MaNGA DR2. For our data reduction, we employed SCRAPPy (section 3), which utilises the `VORBIN` algorithm for data binning (Cappellari & Copin, 2003) and a combination of `pPXF` (Cappellari, 2017) and custom algorithms for the recovery of the stellar kinematics, stellar continuum and emission line properties, making use of the full MILES empirical stellar library, consisting of 985 stellar templates (Sánchez-Blázquez et al. 2006). Unfortunately, this library only covers the wavelength range of 3525-7500 Å, hence not fully exploiting the full coverage of the MaNGA data, which extends up to 10 000 Å. However, it covers the continuum beneath all nebular lines of interest, namely [OII]$\lambda\lambda$3727,3730; Hβ; [OIII]$\lambda\lambda$4960,5008; [OI]$\lambda\lambda$6300;6366; [NII]$\lambda\lambda$6550,6585; Hα; and [SII]$\lambda\lambda$6718,6733, and has a spectral resolution high enough to match the MaNGA spectra (R$\sim$2000).

Since outflows (both AGN- and SF-driven) are generally more prominent in the central region of galaxies, we have first inspected the spectrum of the central region of the galaxy, i.e. extracted from the central 2.5 arcsec (which corresponds to the typical PSF of the Sloan 2.5m Telescope), with the goal of identifying candidate galaxies showing indication of outflows. One aside to the decision for this scale is that, due to the MaNGA galaxies spanning a range of redshifts, this area from which we extract our 'central' spectrum will correspond to a different physical scale on each galaxy. We placed the default emission line component requirements of SCRAPPy onto out component fits (minimum SNR of 3, minimum EW of 3Å), and if the fit to the central spectrum returns two components for the emission lines, then we define the galaxy as an outflow host candidate.

Next, SCRAPPy estimated the continuum SNR of each spaxel through the default processes within the SDSS r-band. All spaxels with SNR < 5 were removed at this point, as the spaxels were to be binned to higher SNR and these spaxels would lead to large bins with noisy data.

To ensure our computation times were kept to a minimum, we utilised the library-trimming module of SCRAPPy (section 3.4.4). Taking the remaining good spaxels, all of the spectra were collapsed into a single integrated spectrum, which was fitted by SCRAPPy and a bespoke 'trimmed' library was produced, typically reducing the full MILES library from 985 to 25−40 templates that best described the galaxy properties.

Following this, we ran SCRAPPy up to its second full fitting iteration: the spaxel fitting

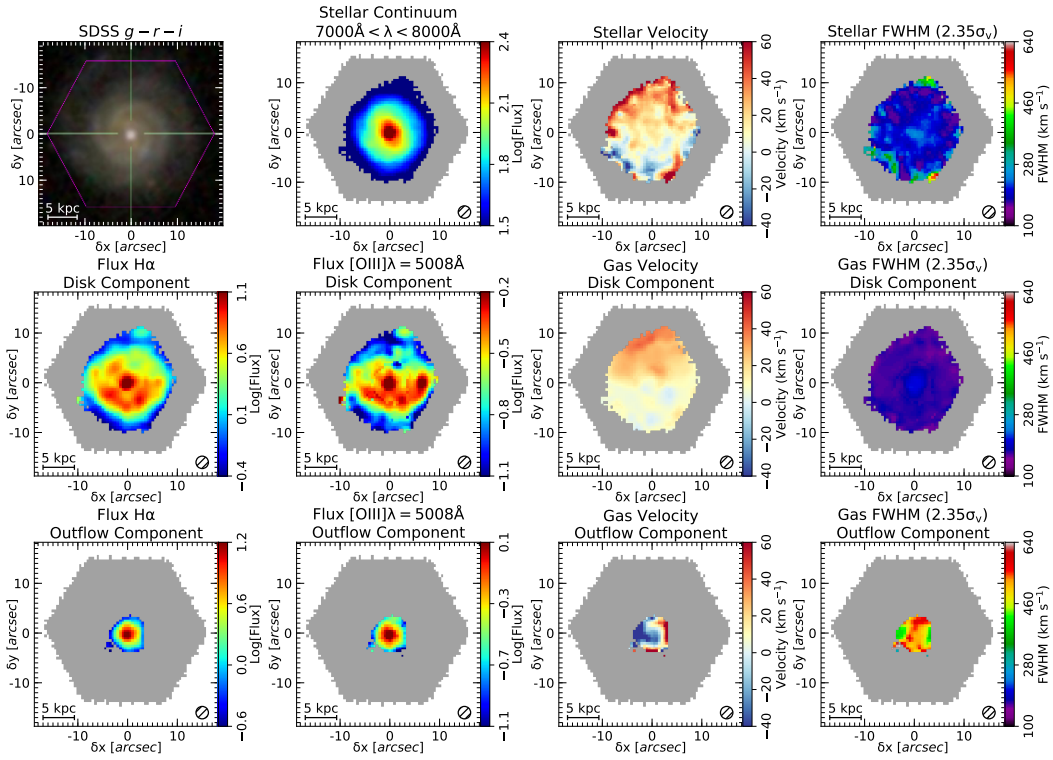


Figure 5.1: Example multi-component mapping. For a representative galaxy the panels show flux, velocity and velocity dispersion maps for the stellar population, for the narrow component of the nebular lines, tracing the galactic disk, and for the broad component of the nebular lines, tracing the outflow. The grey regions indicate spaxels within the MaNGA fibre bundle footprint but masked because either the stellar continuum or emission line component is not detected.

module. Although we applied SCRAPPy to all 2800 galaxies available, our final analysis could only be performed for a total of 37 MaNGA galaxies. This does not mean that other galaxies in the MaNGA data do not show outflows, but that only for these 37 galaxies were all of the lines required for BPT analysis prominent, and that the outflowing gas could be kinematically separated from the constituent gas of the galaxy disk, allowing for a clean and detailed analysis of the outflow using BPT diagnostics.

The list of the 37 galaxies with outflows that can be characterised through the BPT diagnostics is given in table 5.1, along with their stellar masses from the MPA-JHU catalogue (Brinchmann et al. 2004).

Finally, we mention that the fluxes of the nebular emission lines, both narrow and broad, were corrected for dust reddening by assuming the default Calzetti et al. (2000) attenuation curve and using the Hα/Hβ ratio to infer reddening.

5.1.2 Investigating the effects of beam smearing

A potential problem of the modest angular resolution of the MaNGA data is that the putative broad component in the innermost, central region may simply be an artefact of beam smearing of the rotation curve in the central region. Many authors explore this issue by modelling the rotation curve and then simulating the effect of the beam smearing; however, such a test is dependent on the model adopted to describe the velocity field in the central region; therefore, to avoid being model dependent, we have not used this method. Instead, we have simply compared the velocity dispersion of the broad component in the central region with the central stellar velocity dispersion, which is affected by the same beam smearing. A broadening resulting from beam smearing of the stellar velocity field should result in the stellar velocity dispersion being as high as the broad component of the gas or potentially even larger, given that the central region stars are also supported by velocity dispersion. Figure 5.2 shows comparisons of the central velocity dispersion of both the narrow (left) and broad (right) gas components with the stellar velocity dispersion. The narrow components tend to show a similar dispersion as the stars (the solid line represents the 1:1 ratio), and in some cases a smaller dispersion, as the gaseous disk is typically dynamically cooler than the stars in the bulge (e.g. Bertola et al. 1995; Corsini et al. 1999; Young, Bureau & Cappellari 2008; Martinsson et al. 2013). Conversely, the broad components are always found to have a dispersion much larger than the stellar component (the dashed line represents the 2:1 ratio), indicating that the broadening cannot originate from beam smearing. This further indicates that the broad component is tracing gas not in virial equilibrium with the gravitational field, which is traced by the old stellar population, hence implying that the broad component is truly tracing outflowing material. A further indication that the broad component is not resulting from effects of beam smearing is the fact that such a component is not distributed along the minor axis of rotation (which instead is a typical signature of broadening by beam smearing of the rotation field), whilst it is well extended and fully resolved along the major axis direction. The fact that the outflow is resolved is further confirmed by the fact that the broad components have clear kinematic structures both in the velocity map and in the dispersion map.

5.2 Fraction of outflows hosting SF

5.2.1 Spatially resolved BPT-diagnostics

We have explored the properties of the outflow by investigating the spatially resolved BPT diagrams of the outflow in each galaxy.

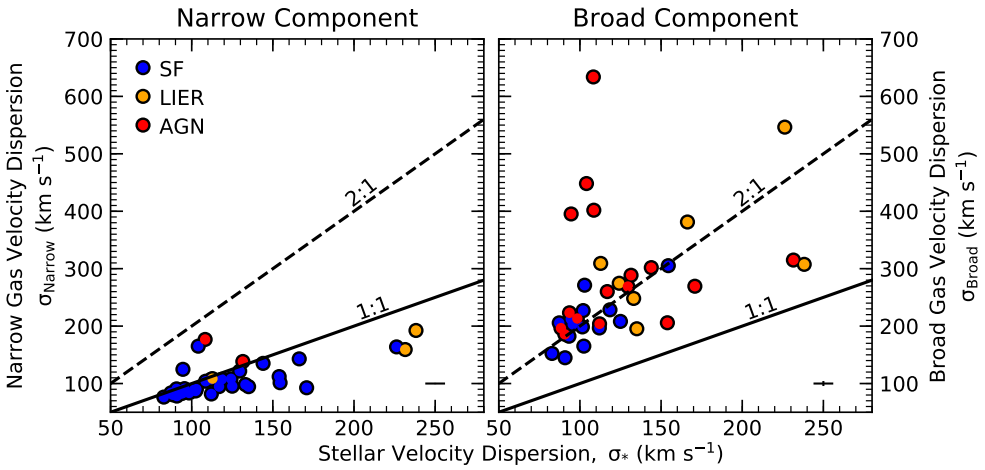


Figure 5.2: Left: Velocity dispersion of the narrow component of the nebular lines in the central region compared with the stellar velocity dispersion. As expected, the narrow component has a velocity dispersion similar to the stellar velocity dispersion, or even lower, owing to the fact that the gas disk is generally dynamically colder than stars. Right: Velocity dispersion of the broad component of the nebular lines in the central region compared with the stellar velocity dispersion in the same region. The former is much larger than the latter, indicating that the broad component cannot result from beam smearing effects of the central rotation curve and that it must be associated with non-virial motions, i.e. outflows. The points are colour-coded by their median [SII]-BPT classifications, as shown in figure 5.4

Figure. 5.3(a) shows the lines of demarcation between the different excitation mechanisms identified by Kauffmann et al. (2003), Kewley et al. (2001), and Kewley et al. (2006) in the three BPT diagrams. The grey shaded contours show the distribution of several hundred thousands galaxies in the SDSS (single fibre) DR7 survey. The distribution of spatially resolved MaNGA spaxels for the full survey is broadly consistent with the single-fibre DR7 data; however, we prefer to use the latter as the MaNGA spaxels underpopulate the AGN locus (Belfiore et al. 2016).

We note that other diagnostic diagrams involving the equivalent width of Hα (Cid Fernandes et al. 2011) cannot be used in this case, as we cannot discriminate the continuum associated with the broad and narrow components. Therefore, in this section, we only exploit the BPT diagrams to identify the excitation mechanism in galactic regions and specifically in the outflows.

As an example of BPT diagnostics spatially resolved in outflows, figure 5.3(a) shows the BPT classification of the broad (outflow) component of each individual spaxel of the same representative galaxy shown in figure 5.1. The representative error bars shown in the bottom right show the median errors for each of the ratios (where the errors are determined from the errors returned on the fit to the velocities, velocity dispersions, and fluxes for each line). Figure 5.3(b) shows how the BPT classification is distributed within the outflow map. For this galaxy,

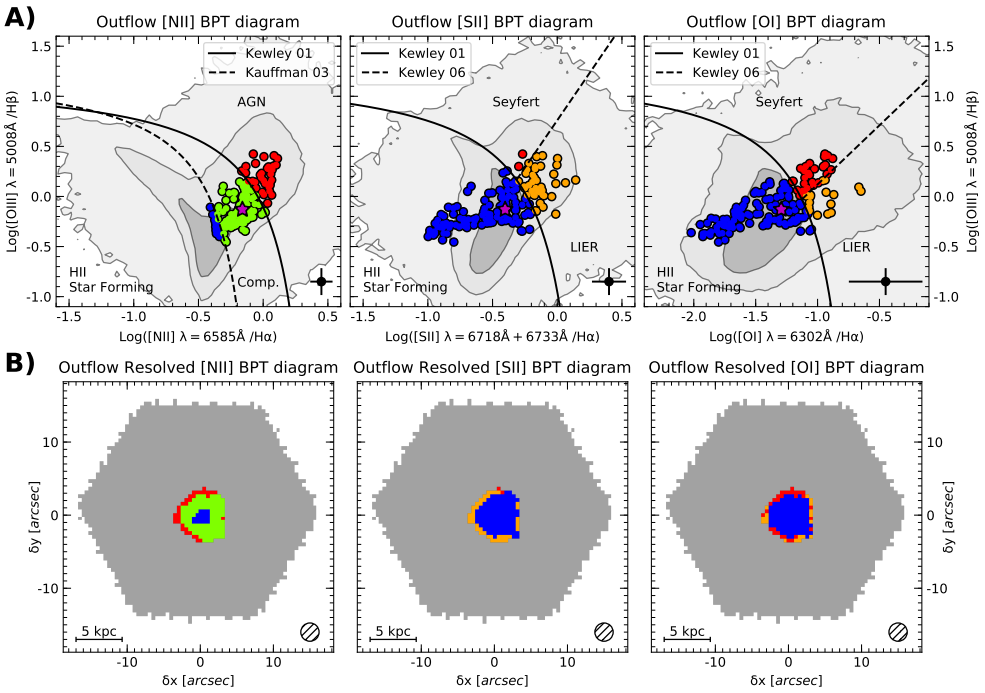


Figure 5.3: (A): Distribution of the broad (outflow) component of the nebular lines on the BPT diagrams for the same representative galaxy of figure 5.1. A significant fraction of this galactic outflow is in the BPT regions typically populated by star forming regions and SF galaxies. The magenta star shows the median location of the BPT points in the outflow. The background grey-scale contours indicate the distribution of several hundred thousand galaxies in the SDSS survey. In all panels, the blue points indicate spaxels classified as SF. In the [NII]-BPT diagram, the green points indicate spaxels classified as Composite, whilst the red points indicate spaxels classified as AGN. On both the [SII]- and [OI]-BPT diagrams, the orange points indicate spaxels classified as LIER, whilst the red points indicate spaxels classified as Seyfert. (B): Spatially resolved BPT classification of the outflowing gas, using the same colour coding as in (A). The central region of the outflow is mostly SF, while towards the outer parts of the outflow excitation by shocks likely dominates.

both the [SII]-BPT and [OI]-BPT diagrams consistently indicate that the central region of the outflow is SF dominated, whilst shocks likely dominate in the outer regions of the outflow. The [NII]-BPT diagram also provides a consistent picture if one includes 'composite' spaxels (green) as hosting SF, since, as discussed above, objects in this region of the BPT diagram can still potentially be fully SF. Once again, one should also take into account that the [NII]-BPT diagram is considered as the least reliable of such diagrams because of its strong dependence on the nitrogen abundance (Masters, Faisst & Capak 2016) and on the ionisation parameter (Strom et al. 2018).

To have a useful quantity of comparison for the outflow in the different galaxies in our

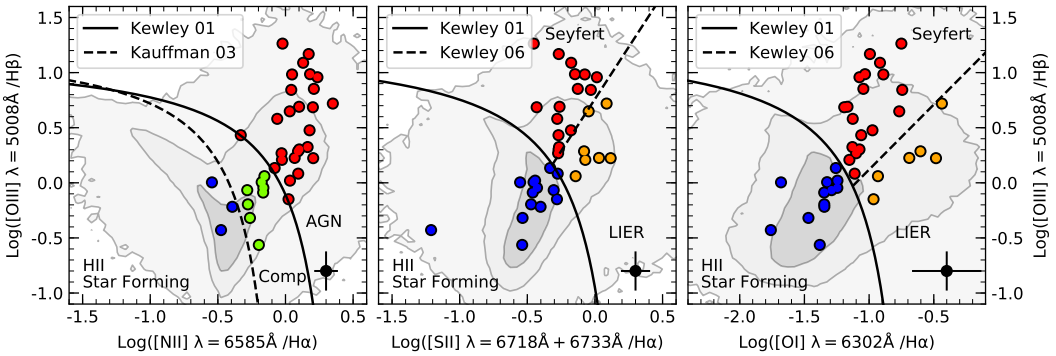


Figure 5.4: Median BPT classification of the outflows for the galaxies in our sample. According to their median BPT classification in the [SII] and [OI] diagrams (central and right panes), about 30% of the galactic outflows in our sample are 'star forming'. The same applies to the [NII] classification if one includes galaxies classified as 'composite' (however, one should take into account that the [NII] classification is more ambiguous, as discussed in the text).

sample, we condense the information on each outflow by taking the median location on the BPT diagram of such outflow components for each galactic outflow. In the case of the galaxy shown in figure 5.3(a), the median location in each BPT diagram is shown with a magenta star symbol. We note that this is a considerable improvement, enabled by spatially resolved spectroscopy, relative to spatially integrated spectra, which are generally light-weighted and therefore dominated by the brightest regions (e.g. a central AGN-dominated region).

Figure 5.4 shows the resulting median location of the outflows on the BPT diagrams for all systems in our sample (Table 5.2 gives the average BPT classification for each individual object). The [SII]-BPT and [OI]-BPT diagrams indicate that in most outflows the ionised gas appears to be excited by AGN; however, about 30 per cent of the outflows host gas whose ionisation is dominated by young stars. For the [NII]-BPT, a smaller fraction of outflows appear dominated by SF excitation. However, if one also includes outflows whose [NII]-BPT classification is in the 'composite' region (considered as a transition region in which SF and AGN excitation coexists), then the fraction of outflows showing indication of SF increases to about 30 per cent also in the [NII]-BPT diagram. One should also take into account the concerns about the [NII]-BPT diagram discussed previously.

Even for those galactic outflows whose median BPT diagnostics are associated with AGN or LIER-like ionisation, there are generally still regions of the outflow that are associated with SF. Therefore, to better quantify the occurrence of SF in galactic outflows, we have estimated for each outflow the fraction of spaxels that can be classified as 'SF' according to each of the BPT diagnostics. Figure 5.5 shows the distribution of galaxies as a function of the fraction of the outflow that is classified as star forming. According to the [SII]-BPT and [OI]-BPT

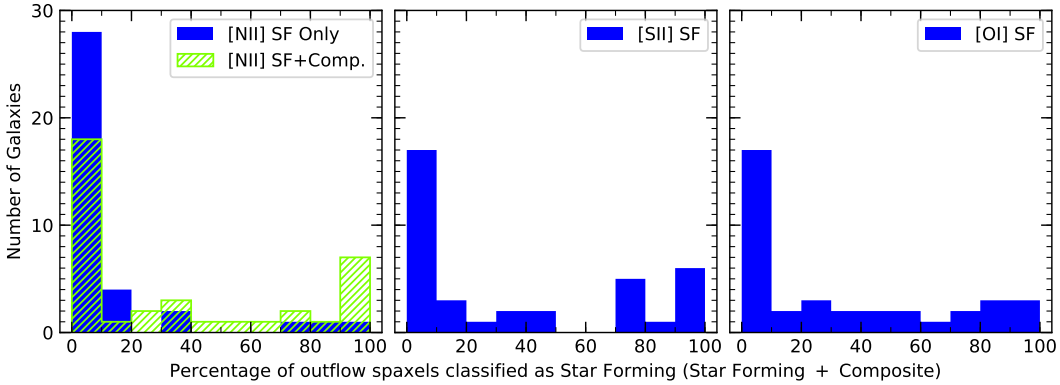


Figure 5.5: Distribution of the outflows as a function of their SF fraction. Each panel shows the distribution of outflows which have a given fraction of spaxels classified as SF according to each of the three BPT diagrams. In the case of the [NII]-BPT the combined case of 'SF' and 'composite' is also shown. According to the [SII] and [OI] classifications about half of the outflows have at least 10% of the spaxels that are SF. The same result applies to the [NII] classification if one includes the composite cases.

classifications, about half of the galactic outflows show evidence for some star formation inside the outflow (whose contribution to the overall excitation of the gas ranges from 10 per cent to 100 per cent). Even the [NII]-BPT diagram provides the same fraction of 'star-forming outflows' if one considers also outflowing regions classified as 'composite' in this diagram.

5.2.2 *In-situ* versus external photoionisation

As discussed in chapter 4, one potential concern of our finding could be that, for those outflowing regions where the BPT diagnostics of the outflowing gas are consistent with excitation from young stars, the latter may not be located inside the outflow but in the galactic disk, and the outflowing gas may simply be illuminated externally by the UV radiation coming from the disk. However, these scenarios can be distinguished by exploiting the ionisation parameter, as pointed out in chapter 4

Figure 5.6 shows the [OIII]λ5008/[OII]λ3727 ratio as a function of the R23 parameter. The right-hand axis translates the [OIII]λ5008/[OII]λ3727 ratio into U following the relation provided by Díaz et al. (2000). The shaded contours indicate the distribution of star-forming galaxies from the SDSS survey (68 per cent, 95 per cent, and 99.7 per cent of the population), illustrating the well-known correlation between U and metallicity (e.g. Nagao et al. 2006). Dark blue symbols indicate the median values observed in the galactic outflows of our sample that are dominated by star formation, according to the [SII]-BPT diagnostic diagram, as in figure 5.4, whilst light blue symbols show the medial location of the narrow component tracing SF in

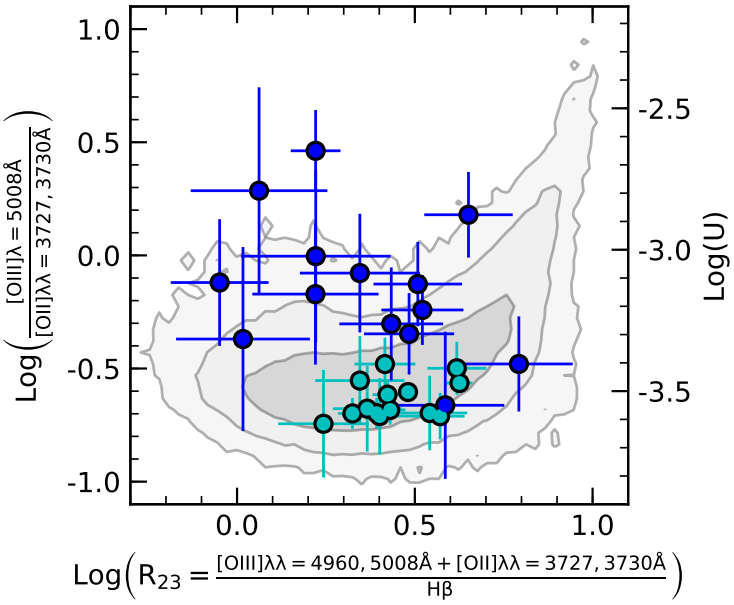


Figure 5.6: Ionization parameter (right hand y-axis) as traced by the [OIII]/[OII] ratio (left hand y-axis) as a function of the R_{23} parameter, which is primarily sensitive to the gas metallicity. Dark blue symbols indicate the median values observed in those galactic outflows dominated by SF, according to the[SII]-BPT diagnostic diagram, as in figure 5.4. The light blue symbols are the average location of the narrow component in regions that are classified as SF. The shaded contours indicate the distribution of SF galaxies from the SDSS survey (68%, 95% and 99.7% of the population). SF outflows (blue symbols) have similar ionisation parameter as normal SF regions, and possibly even slightly higher. The error bars shown are the median error of the ratios calculated for all spaxels within the outflow.

the galaxy disk. Figure 5.6 shows that the ionisation parameter of the gas in the starforming outflows, as inferred from the [OIII]λ5008/[OII]λ3727 line ratio, is undistinguishable from that of normal star-forming regions/galaxies and, if anything, is even slightly higher. Given that the density in the outflows is higher than that in the disk (as explored in chapter 6), this result excludes that the photoionisation in these outflows is due to external UV radiation coming from the star-forming disk, and confirming that the outflowing gas is excited by *in-situ* (i.e. within the outflow) SF.

5.2.3 Contribution by shocks

One caveat to using both the BPT diagrams and the R23 versus O32 diagram in identifying *in-situ* SF lies in the fact that, although generally shocks populate the LIER-like region of these diagram (e.g. Allen et al. 2008; Diniz et al. 2017), some peculiar shocks (especially low velocity ones) can potentially produce star-formation-like emission line ratios on the BPT

diagrams. In chapter 4, this form of potential shock excitation was excluded through the use of IR diagnostics, namely the [FeII]λ1.64μm/Brγ versus H_2(1−0)S(1)/Brγ and [Fe II]λ1.25μm/Paβ versus [PII]λ1.18μm/Paβ diagrams, afforded by the extended IR wavelength coverage offered by the X-shooter data. Unfortunately, the MaNGA wavelength coverage stops at ∼10 000Å, hence preventing their use for disentangling SF in these galaxies.

However, we have investigated whether such shocks can potentially mimic SF-like line ratios observed by us in galactic outflows by employing the MAPPINGS III library of fast radiative shock models (Allen et al. 2008), using models that span a broad range of velocities (v_s = 100−1000 km s^{-1}) and magnetic parameters ($B/\sqrt{n}$ = 10^{-4} − 10μGcm$^{3/2}$). We have adopted solar abundances (appropriate for the central region of these galaxies; Belfiore et al. 2016) and pre-shock densities of 1 cm^{-3} (which is the case most extensively studied by Allen et al. 2008 and generally regarded as appropriate for the pre-shock conditions). The predictions of the extensive grids of these models, spanning a wide range of velocities and physical parameters, have been overlayed on the BPT diagrams and on the R23 versus O32 diagram, together with the location of the line ratios observed in the MaNGA outflows in figures 5.7 through 5.10. Not unexpectedly, shock models can reproduce well several line ratios in the LIER region, as well as in the AGN-Seyfert regions. However, although some shock models do overlap with the loci occupied by star-forming regions, the vast majority of the line ratios observed in the star-forming outflows identified by us (blue symbols) are inconsistent with the shock models, further reinforcing the claim that we are indeed observing SF inside these outflows.

The additional issue with shocks is that they are much less effective in producing ionising photons than young, hot stars, hence much less likely to account for the large nebular luminosities observed in the outflow. Assuming a typical radius of the putative shocked region of ∼3 kpc (as inferred from our maps), by using the equations given in Allen et al. (2008), focusing on slow shocks (v_s ∼ 100 km s^{-1}), which are the ones that may mimic star-forming like line ratios, and assuming the extreme (unrealistic) case of shocked gas covering the entire 4π solid angle with covering factor of unity, we infer that the expected ionising photon luminosity should be about $Q_{ion} \approx 10^{52}$ s^{-1}. Assuming that these photons are entirely absorbed by the gas (by ionising it), this would result into a Hα luminosity of the outflowing gas of ∼10^{40} erg s^{-1}. Much more realistically, the putative shocked gas is characterised by a much smaller solid angle (outflows, when observed at high angular resolution have typically solid angles less than π (e.g. Venturi et al. 2018), with small covering factor (0.02−0.2, e.g. Baskin & Laor 2005) and only a fraction of the ionising photons is likely absorbed by the outflowing clouds, hence resulting in to an expected Hα luminosity one or two orders of magnitude lower, i.e. ∼10^{38}−10^{39} erg s^{-1}. Except for one case, all outflows that we have classified as SF have a Hα luminosity in the range of 10^{40}−10^{41} erg s^{-1}, hence are very unlikely to be produced by shocks.

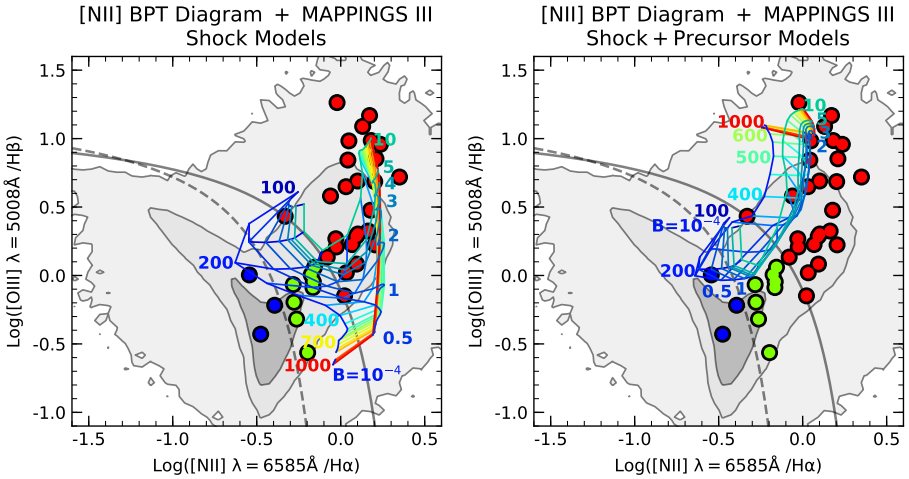


Figure 5.7: [NII]-BPT diagram showing the median location of the outflowing gas in each of the galaxies our our sample, as figure 5.4, but also showing the result of a grid shock models with varying velocities, from 100 km s^{-1} (blue) to 1,000 km s^{-1} (red), and varying magnetic field parameter, from B/n$^{1/2}$ = 10^{-4} μG cm$^{3/2}$ (dark blue) to B/n$^{1/2}$ = 10 μG cm$^{3/2}$ (light blue), both without (left) and with (right) shock precursor. Outflows classified as SF can hardly be accounted for through shock models.

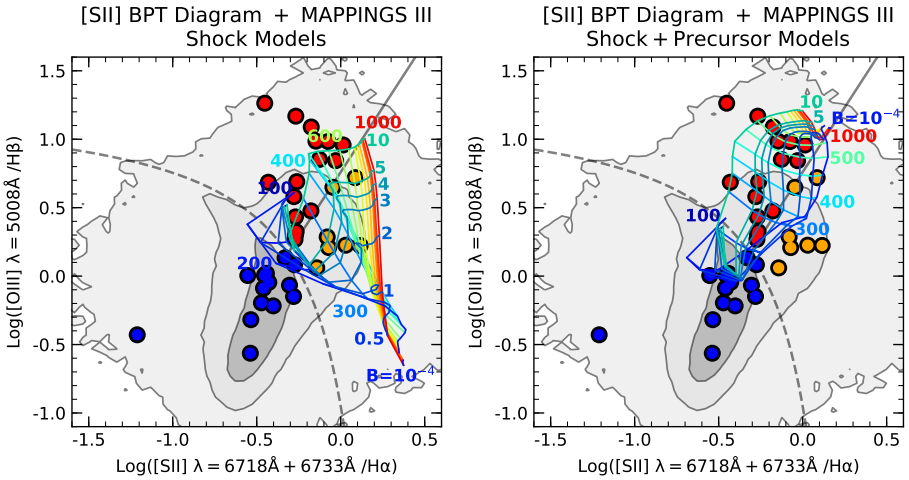


Figure 5.8: [SII]-BPT diagram showing the median location of the outflowing gas in each of the galaxies our our sample, as figure 5.4, but also showing the result of a grid shock models with varying velocities, from 100 km s^{-1} (blue) to 1,000 km s^{-1} (red), and varying magnetic field parameter, from B/n$^{1/2}$ = 10^{-4} μG cm$^{3/2}$ (dark blue) to B/n$^{1/2}$ = 10 μG cm$^{3/2}$ (light blue), both without (left) and with (right) shock precursor. Most outflows classified as SF cannot be explained in terms of shocks.

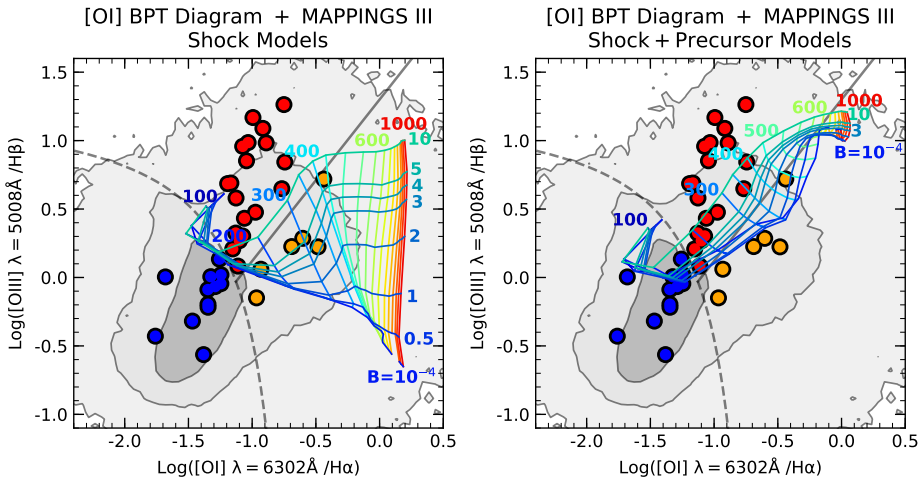


Figure 5.9: [OI]-BPT diagram showing the median location of the outflowing gas in each of the galaxies our our sample, as figure 5.4, but also showing the result of a grid shock models with varying velocities, from 100 km s^{-1} (blue) to 1,000 km s^{-1} (red), and varying magnetic field parameter, from B/n$^{1/2}$ = 10^{-4} μG cm$^{3/2}$ (dark blue) to B/n$^{1/2}$ = 10 μG cm$^{3/2}$ (light blue), both without (left) and with (right) shock precursor. Most outflows classified as SF cannot be explained in terms of shocks.

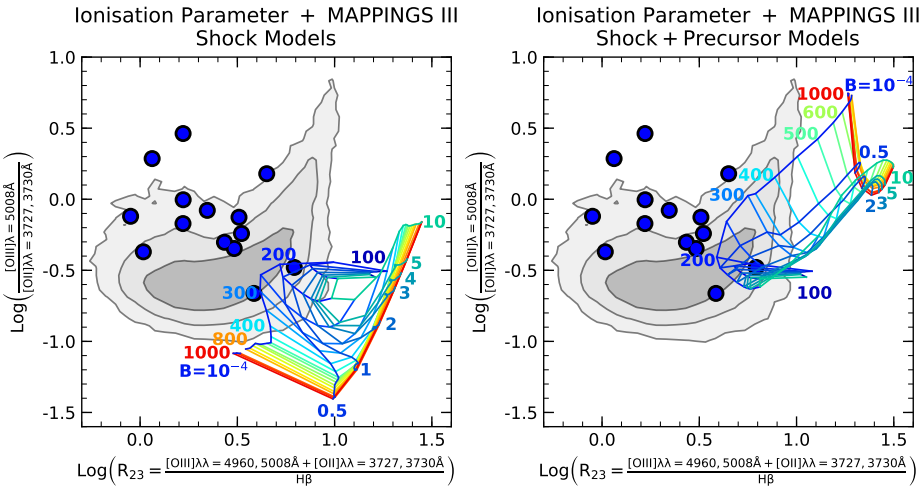


Figure 5.10: R_{23} vs [OIII]/[OII] diagram showing the median location of the outflowing gas in each of the galaxies our our sample, as figure 5.6, but also showing the result of a grid shock models with varying velocities, from 100 km s^{-1} (blue) to 1,000 km s^{-1} (red), and varying magnetic field parameter, from B/n$^{1/2}$ = 10^{-4} μG cm$^{3/2}$ (dark blue) to B/n$^{1/2}$ = 10 μG cm$^{3/2}$ (light blue), both without (left) and with (right) shock precursor. Most outflows classified as SF cannot be explained in terms of shocks.

The combination of these energetics arguments and the location of shock models on the diagnostic diagrams discussed above provide strong evidence against any significant role of shocks in the ionisation of the outflows that we have identified as hosting SF

5.3 SFR inside outflows

5.3.1 Contribution to the total SF

For those outflows showing evidence for SF, we can estimate the SFR by using the broad component of Hα as a tracer of SF (Kennicutt & Evans 2012). The SFRs inferred in these local outflows are of the order of 0.1-1 $M_{\odot}$ yr^{-1} (Table 5.2). In figure 5.11(a), we compare the SFR inside the outflow with the total SFR of the galaxy, as inferred by the total Hα emission (both narrow and broad), integrated across the entire galaxy (excluding AGN and LIER-like regions). SF in the outflow contributes 'only' between 5 per cent and 30 per cent to the total SFR of these local galaxies (with one extreme case dominating the total SFR).

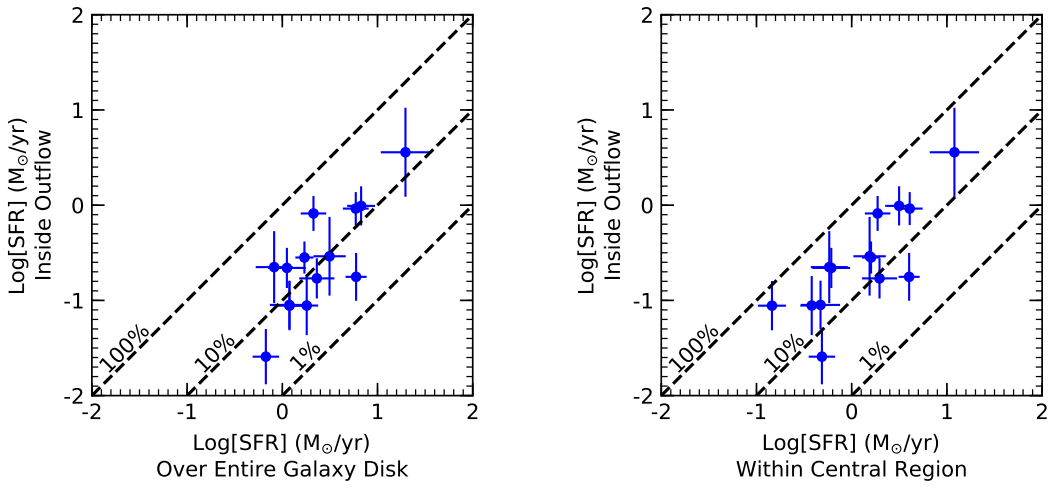
However, in the spaxels where multiple components are detected (typically within the central few kpc), the SF inside the outflow can dominate the local SFR. This is illustrated in 5.11(b) where the SFR inside the outflow is compared with the SFR in the region in which the outflow is detected (i.e. where a broad component is detected, which is generally in the central region).

These SFRs are certainly modest and may be regarded as currently not very relevant for galaxy evolution. However, one should also take into account that these are relatively mild outflows, whilst SF may be much more prominent inside the massive outflows driven by massive star-forming galaxies at high redshifts or powerful distant quasars. To investigate this possibility in the following, we explore the scaling relation between SF inside the outflow and mass outflow rate.

5.3.2 Ionised outflow rate

Unfortunately, we only have information on the ionised component of the outflowing gas, therefore we can only obtain information on the ionised gas outflow rate. The ionised phase often accounts for a small fraction of the total gas content in outflows (Carniani et al. 2015; Fiore et al. 2017; Fluetsch et al. 2019); however, it can be used as a proxy of the global outflow rate.

To calculate our outflow rates we utilise the equations and methods laid out in section 1.3.2. Specifically, we estimate the contribution to the outflow rate for each spaxel using equation 1.3, defining the radius as the distance of the spaxel from the outflow centre, and the velocity



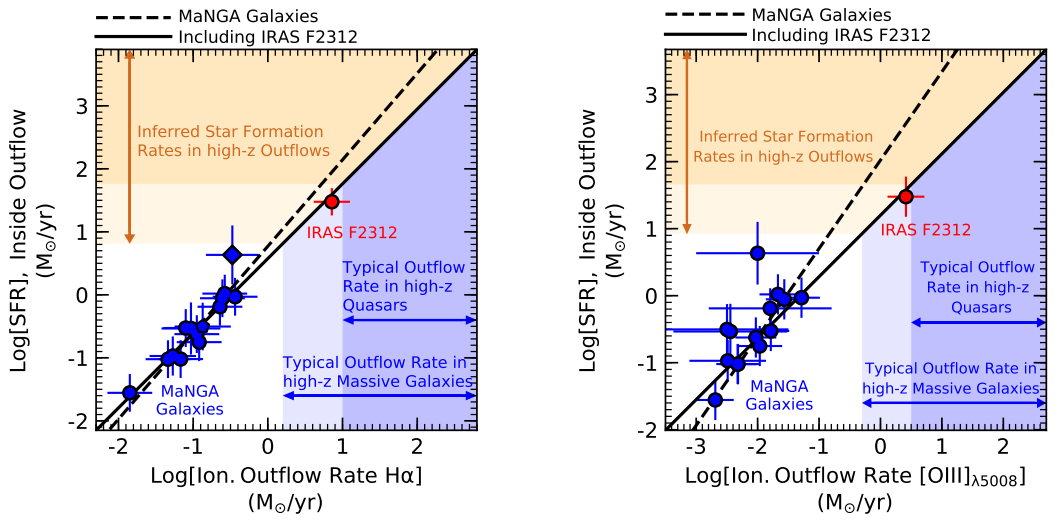
(a) (b)

Figure 5.11: (a) . SF rate inside the outflow as a function of total SF rate in the disk of the host galaxy. (b) SF rate inside the outflow as a function of the SF rate in the same central projected area in which the outflow is detected.

in that location as $|v| + FWHM/2$ of the broad component ($|v|$ being the absolute velocity of the centre of the broad component tracing the outflow, relative to the galaxy rest frame). We determine the outflow mass in each spaxel using equations the [OIII] and Hα luminosities of the broad component in equations 1.6 and 1.7, assuming a density of 200 cm^{-3}, which is intermediate between the results obtained by Perna et al. (2017) (for AGN-driven outflows) and the result obtained by Genzel et al. (2014) (for outflows in massive star-forming galaxies). In chapter 6, we refine the outflow rates by using the average outflow gas density inferred from the sulphur doublet for a large sample of outflows, and develop a more realistic method to determine outflow rates. In both cases, the total outflow rate was obtained by summing the individual contributions to the outflow rate of all individual spaxels.

5.3.3 Scaling relations and implications for high-redshift galaxies

We have explored the connection between outflow rate and star formation in the outflow. We have initially focused on the outflow rate inferred from the Hα broad component, as it is less subject to ionisation and metallicity corrections. For those galaxies whose outflow is dominated by SF in terms of BPT diagnostics, figure 5.12(a) shows the SFR inside the outflow as a function of the ionised outflow rate estimated from the broad component of Hα. Blue symbols are star-forming outflows discovered in this paper, whilst the red symbol is the star formation previously discovered inside the outflow of the galaxy IRAS F23128-5919 (chapter



(a) (b)

Figure 5.12: (a): SF rate inside the outflow as a function of ionised gas mass outflow rate inferred from the Hα. Blue symbols are SF outflows in local MaNGA galaxies studied in this paper. The red symbol indicates the SF outflow previously detected in IRAS F23128-5919 (chapter 4). The dashed line shows the fit to the MaNGA outflows presented in this paper (blue points). The solid line is the linear fit including IRAS F23128-5919 (equation 5.2). The blue shaded regions indicate the ionized outflow rate inferred for high-z massive galaxies and quasars. The orange shaded regions indicates the SF rate inside these distant galactic outflows if they follow the same correlation observed locally. (b): as panel (a), but where the ionised outflow rate is now inferred from the broad component of the [OIII] line.

4). The data reveal a clear correlation between the two quantities. To perform linear fits to the data, accounting for the errors in both the outflow rate and SFR, we used the `scipy` Orthogonal Distance Regression (ODR) package. The errors in these quantities were calculated through rigorous error propagation, though we elected to maintain a minimum error of 0.3 dex to account for previous considerations, such as assuming density values of 200 cm^{-3} and LOSVD errors being the formal errors on the fit returned from `mpfit`. We began by performing a fit to the MaNGA outflows presented in this paper alone. This fit is shown by the dashed line, and extrapolation of this fit is nicely consistent with the SF observed in the outflow of IRAS F23128-5919. This fit takes the form:

$$\log(\mathrm{SFR_{outfl.}}\ [\mathrm{M_\odot yr^{-1}}]) = 1.37(\pm0.14)\ \log(\dot{M}_{\mathrm{outfl-ion,H}\alpha}\ [\mathrm{M_\odot yr^{-1}}]) + 0.77(\pm0.14) \quad (5.1)$$

Including IRAS F23128−5919 in the fit results in the fit shown by the solid line, which is only slightly different from the fit to the MaNGA galaxies alone, and takes the form:

$$\log(\mathrm{SFR_{outfl.}}\ [\mathrm{M_\odot yr^{-1}}]) = 1.19(\pm0.07)\ \log(\dot{M}_{\mathrm{outfl-ion,H}\alpha}\ [\mathrm{M_\odot yr^{-1}}]) + 0.58(\pm0.07) \quad (5.2)$$

We recall again that generally the ionised outflow rate is a small fraction of the total outflow rate (especially relative to the molecular phase). The blue shaded regions indicate the ionised outflow rate inferred for high-z massive galaxies and quasars, as inferred from Hα or Hβ data (Carniani et al. 2015; Genzel et al. 2014; Fiore et al. 2017). The orange shaded regions indicate the SFR inside these distant galactic outflows if they follow the same correlation observed locally, implying SFRs in outflows ranging from $\sim 10\ \mathrm{M_\odot\ yr^{-1}}$ to $\sim 1000\ \mathrm{M_\odot\ yr^{-1}}$, which would certainly contribute significantly to the galaxy formation, and to their spheroidal component in particular. Such high SFRs in distant galactic outflows may have been missed by past observing campaigns, and in section 5.5, we will discuss the possible reasons for this.

One potential problem of the diagram in figure 5.12(a) is that the broad Hα flux is a quantity used to measure both the ionised outflow rate (together with the additional information on the outflow velocity and distribution) and to estimate the SFR in the outflow, which may induce a spurious correlation. Therefore, in order to test whether the correlation persists even when adopting different tracers that use independent quantities, we have considered the outflow rate estimated through the [OIII] line. As discussed above, [OIII] is more unreliable as an outflow tracer (because of its dependence on the metallicity and on the ionisation degree) and tends to underestimate the outflow rate; however, it has the advantage that it is completely independent of the Hα flux used to estimate the SFR. Figure 5.12(b) shows the SF in the outflow as a function of the ionised outflow rate as inferred from the broad component of the [OIII] line. Again, two fits were performed: first to the MaNGA outflows alone, and secondly to the MaNGA outflows including IRAS F23128-5919 in the fit. The two quantities are still clearly correlated, although the scatter is slightly larger than in figure 5.12(a). In this case, the fit to the MaNGA outflows alone (dashed line) is less consistent with the outflow in IRAS F23128-5919, and takes the form:

$$\log(\mathrm{SFR_{outfl.}}\ [\mathrm{M_\odot yr^{-1}}]) = 1.32(\pm0.18)\ \log(\dot{M}_{\mathrm{outfl-ion,[OIII]}}\ [\mathrm{M_\odot yr^{-1}}]) + 2.02(\pm0.37) \quad (5.3)$$

However, the larger scatter in the data makes the extrapolation of the MaNGA data less reliable. The solid line shows the linear fit including both the MaNGA outflows and the outflow in IRAS F23128-5919, which is given by the following equation:

$$\log(\mathrm{SFR_{outfl.}}\ [\mathrm{M_\odot yr^{-1}}]) = 0.92(\pm0.06)\ \log(\dot{M}_{\mathrm{outfl-ion,[OIII]}}\ [\mathrm{M_\odot yr^{-1}}]) + 1.20(\pm0.10) \quad (5.4)$$

The slope, close to unity, is very similar to the slope obtained in equation 5.2 when using Hα as a tracer of the SFR. The intercept is obviously different, as a consequence of [OIII] underestimating the ionised outflow rate. The blue and orange shaded regions in figure 5.12(b) have the same meaning as in figure 5.12(a) (in this case the range of outflow rates, based on [OIII], in high-z galaxies is more incomplete and we have simply assumed the ranges inferred from Hα and Hβ, and shifted by a factor of 0.5 dex, as discussed in section 5.3.2). As in the previous case, the extrapolation to the outflow rates of high-redshift galaxies implies high SFRs in the outflows of distant galaxies.

We shall recall that stars obviously form out of the molecular phase of galactic outflows; therefore, it would certainly be more appropriate to exploit scaling relations involving the molecular outflow rate. Unfortunately, information on the molecular outflow rate is not yet available for these star-forming outflows (although ALMA observations are ongoing for some of these systems). However, it should be noted that Fluetsch et al. (2019) have found that the molecular-to-ionised outflow rate increases with AGN luminosity (an opposite trend with respect to what found by Fiore et al. 2017, who however use disjoint samples to probe the ionised and molecular phases, hence potentially subject to differential selection effects); therefore, in powerful outflows driven by luminous AGNs the fraction of stars forming in outflows may potentially be even larger than that inferred from the scaling relations with the ionised component of outflows given by equations 5.2 and 5.4.

5.4 Distribution of galaxies with outflows on the SFR-M$_*$ diagram

As discussed in the introductions, the distribution of galaxies on the SFR-M$_*$ diagram is often used to identify their properties. The distribution of the 2800 MaNGA galaxies (our parent sample) on the SFR-M$_*$ diagram is shown with grey contours in figure 5.13 (SFR and M$_*$ are obtained from the MPA-JHU DR7 catalogue, Brinchmann et al. 2004).

The main sequence in this sample is less populated than in other samples (such as, for instance, SDSS-DR7) primarily as a consequence of the MaNGA selection criterion of having a flat distribution in stellar mass. The MaNGA galaxies with outflows investigated in this paper, i.e. those for which the outflow can be studied in terms of BPT diagram, are shown with symbols, which are colour-coded according to the [SII]-BPT classification of the gas in the outflow as in figure 5.4. Most of the galaxies with outflow are distributed around the massive end of the main sequence. Outflows hosting SF (blue symbols) are along the main sequence or slightly above it. There are some outflows amongst passive and green-valley galaxies, but only with AGN and LIER-like excitation. The lack of outflows in low-mass galaxies may primarily result from selection effects, as these tend to be fainter and are likely characterised by milder

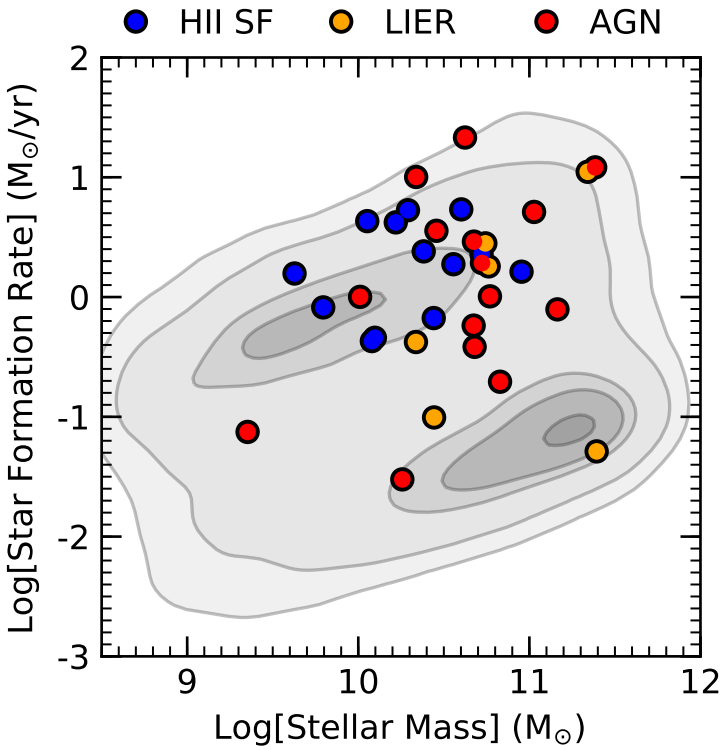


Figure 5.13: SF rate versus stellar mass diagram. The background contour plot shows the distribution of all 2,800 galaxies in the MaNGA sample. The symbols show the distribution of galaxies with outflows presented in this study, i.e. those for which the outflows can be analysed in the BPT diagram. The BPT classification of the outflow is the same as in the [SII]-BPT diagram of figure 5.4. Most outflows, especially those hosting SF inside the outflow, are located around the massive end of the Main Sequence, or slightly above it.

outflows (either SF- or AGN-driven), below our detection threshold. A more detailed analysis of the statistical properties of outflows will be discussed in chapter 6. However, the finding that outflows hosting SF are located primarily on, or slightly above, the massive end of the main sequence supports the scenario in which this SF mode may contribute to the population of stars in the bulge; although the overall contribution to the bulge population at this late epoch is minor, the contribution of this mechanism to the evolution of the bulge may have been significant at high z, in the early phase of galaxy formation. This result is in line with the recent finding that galaxies above the main sequence have bluer bulges, which may still be in the process of forming a fraction of their stars (Morselli et al. 2017). We finally mention that, both morphologically and kinematically, nearly all galaxies showing SF inside outflows appear to be isolated systems, typically regular galaxy disk. Only two of these galaxies are interacting systems, which are marked with superscript 'a' in table 5.1.

5.5 Why has SF in outflows been elusive?

If SF in outflows is as common as inferred in section 5.2 and can reach SFRs as high as inferred in section 5.3.3 in high-z galaxies, why has this phenomenon remained elusive so far?

One of the primary reasons is that different excitation mechanisms do not have the same weight in affecting the location of galaxies or galactic regions on the BPT diagrams. In particular, in this section we show that even if prominent SF is taking place inside the outflow, the presence of an even weak AGN photoionising the gas within the outflow generally dominates these diagnostic diagrams, even in the scenarios with the most conservative assumptions, implying that the fraction of star-forming outflows inferred in section 5.2 is actually a lower limit.

To show this, we make two simplifying assumptions: first, we assume that the BH accretion and SF in the spheroidal component of the galaxy evolve along the local M_{BH}–$\mathrm{M}_{spheroid}$ relation; secondly, we take the most extreme (and most conservative) case possible, wherein we assume that all stars in the spheroid form through SF within the outflow.

This implies that:

$$\dot{M}_{\mathrm{BH}} \approx 5 \times 10^{-3}\ \mathrm{SFR}_{\mathrm{spheroid}} \tag{5.5}$$

(Kormendy & Ho 2013). Assuming a radiative efficiency of 0.1 for BH accretion, the relationship between Hα emission (in the NLR) and X-ray luminosity for type-2 AGNs given in Ueda et al. (2015), and the X-ray to bolometric AGN luminosity correction given in Marconi et al. (2004) (here we take the bolometric correction appropriate for the typical AGN luminosity considered by Ueda et al. (2015) implying $L_{\mathrm{bol}}/L_{(2-10\mathrm{keV})} = 20$), we get:

$$L_{\mathrm{H}\alpha}(\mathrm{AGN})\ [\mathrm{erg\ s^{-1}}] = 2 \times 10^{42} \dot{M}_{\mathrm{BH}}\ [\mathrm{M_{\odot} yr^{-1}}] \tag{5.6}$$

By also using the relation between $L(\mathrm{H}\alpha)$ and SFR given in Kennicutt & Evans (2012), in combination with equations 5.5 and 5.6, we obtain

$$L_{\mathrm{H}\alpha}(\mathrm{AGN}) \approx 5 \times L_{\mathrm{H}\alpha}(SFR) \tag{5.7}$$

and, of course, the same relation applies to Hβ. Moreover, for most local star-forming galaxies typically $L_{[\mathrm{OIII}]}(\mathrm{SFR}) \leq L_{\mathrm{H}\beta}(\mathrm{SFR})$, whilst for AGN $L_{[\mathrm{OIII}]}(\mathrm{AGN}) \approx 10\ L_{\mathrm{H}\beta}(\mathrm{AGN})$; as a consequence, together with equation 5.7, this implies that the BPT diagnostics will be totally skewed towards the AGN region on their Y-axis. Similarly, given that [NII], [SII], and [OI] are all more luminous in AGNs than in SF galaxies, together with equation 5.7, this implies that the BPT diagnostics will be all skewed towards the AGN region also on the X-axis. Therefore,

even in the most conservative scenario assumed above, in which the formation of the spheroidal component of the galaxy results entirely from SF in the outflow, implying prominent SF in the outflow, such star formation may go entirely undetected in the BPT diagrams as the excitation diagnostics are totally dominated by AGN excitation.

This issue has been clearly directly demonstrated in detailed spatially resolved cases in which star-forming clumps have remained elusive because diagnostics have been dominated by the photoionisation of an even weak AGN (Santoro et al. 2016). One consideration that complicates things further is the presence of shocks. In these cases, the detection of SF is even more difficult as they are expected to move the diagnostics towards the LIER regions on the diagrams. Summarising, the lack of SF-like BPT diagnostics in outflows does not necessarily exclude the presence of SF inside the outflow, as such outflows may still host prominent SF but may struggle to be detected as other sources of excitation may be dominating the diagnostics. SF inside outflows becomes easier to detect when the AGN has faded away, and when shocks are weak. The implication in the context of this thesis is that the fraction of outflows with SF may be higher than what estimated in section 5.2.

In integrated, spatially unresolved (or poorly resolved) spectra, the light weighting further enhances the bright, central AGN dominated emission (especially in the case of powerful quasars) relative to weaker star-forming regions. This issue is obviously more problematic at high redshift due to the lower spatial resolution. However, despite these difficulties, current data already show some indications of SF in high-redshift outflows. Genzel et al. (2014) investigated the outflows of five high-z galaxies and two of them have diagnostic ratios consistent with SF inside the outflow. The finding that at high-redshift star-forming regions tend to shift towards the AGN locus on the BPT diagrams (e.g. Strom et al. 2017) is a further complication that may have led the classification of some outflows as AGN dominated, whilst their BPT ratios may actually be associated with SF inside the outflow.

Another problem is that often only the [NII]-BPT diagram is used to classify outflows. As discussed in Strom et al. (2018) and Masters et al. (2016), the [NII]-BPT is much more subject to changes in the ionisation parameters and nitrogen abundance relative to the other two BPT diagrams. In section 5.2, we have seen that star-forming outflows are clearly identified in the [SII]- and [OI]- BPT diagrams, whilst in the [NII]-BPT diagram they are located primarily in the 'composite' region. Therefore, studies based only on the [NII]-BPT diagnostics may have misidentified the excitation of galactic outflows. For instance, Leung et al. (2017) classify the outflows of 13 AGNs at $z \sim 2$ only based on the [NII]-BPT diagram; about half of these outflows are located in the 'composite' region of the [NII]-BPT diagram; hence, these outflows could be hosting SF, similarly to the outflows investigated in this paper. Another problem is that in many spectroscopic surveys outflows are searched through blue wings of the [OIII]5008

line. Because of the dominance of this line in AGN-driven outflows, as discussed above, many of these surveys have certainly biased their search towards outflows whose diagnostics are AGN dominated.

Finally, with the growing interest in studying local galaxy outflows in much greater details thanks to improved, high sensitivity and wider field of view integral field spectrometers (e.g. MUSE at the VLT), alongside evidence for SF being found in additional individual outflows studied in detail (Maiolino et al. 2017; Venturi et al. 2018; Fluetsch et al. 2019) we expect to see many highly detailed and successful studies of star formation within outflows within the years to come.

5.6 Implications for galaxy evolution

Stars formed in galactic outflows have completely different dynamical properties relative to the stars forming in galactic discs. Indeed, as soon as stars form in the outflow they only respond to gravity, hence moving ballistically. If the velocity at the time of formation inside the outflow exceeds the escape velocity then the newly formed stars will leave the galaxy and disperse in the IGM. If the velocity is lower than the escape velocity then the newly formed stars will be bound; they will be rapidly decelerated and will start oscillating around the galactic centre. The latter is also one of the reasons why it is difficult to identify the young stars formed in the outflows, since as soon as they form they decouple from the outflow and their velocity is quickly reduced (within a few Myr, chapter 4) becoming kinematically difficult to distinguish from the stars in the disk.

In the case of the stars formed in the outflows that are gravitationally bound, their resulting orbits may predominantly be radial. However, as shown in chapter 4 for the specific case of IRAS F23128-5919, the gravitational interaction with the galactic disk, or any other non-spherically symmetric potential (stellar bar, flyby galaxies) will introduce a tangential component to the stellar velocities, hence randomising their orbits. Moreover if the expelled gas was initially in rotation around the galactic centre, then the resulting stars formed in the outflow star can also preserve some minor rotation component.

What the primary fate is of stars formed in the outflow has still to be investigated properly in a statistical way. Models expect both scenarios, i.e. escaping and gravitationally bound, depending on the nature and properties of the outflow (Zubovas et al. 2013c). In the case of the star-forming outflow in IRAS F23128-5919, which has been studied in great detail, we showed in chapter 4 that most of the stars born in the outflow are gravitationally bound. A detailed study of the MaNGA galaxies is not possible, however regardless of the fate of stars formed in outflows he fact that SF in outflows has been found by us to be so common implies that

this newly discovered phenomenon may have major implications for galaxy evolution and even result into a paradigm change in some specific areas. In the following, we discuss some of the potential implications of star formation in outflows for galaxy evolution and to explain some of the galactic properties.

5.6.1 Formation of the spheroidal component of galaxies

If stars formed in outflows are gravitationally bound (this seems to be the case in most cases, as the broad component velocity dispersion is generally comparable or slightly lower than the escape velocity, as expected by some models, Zubovas et al. 2013a), then they can contribute significantly to the formation and evolution of the spheroidal component of galaxies, i.e. bulge, halo, and even contribute to the formation of elliptical galaxies. Indeed, Silk (2013) have proposed that such positive feedback mechanism may explain some of the extreme star-forming galaxies at high redshift, which are thought to be progenitors of local ellipticals. Ishibashi et al. (2013) have even suggested that SF in AGN-driven outflows can contribute to the observed size evolution of elliptical galaxies across the cosmic epochs. Wang & Loeb (2018) have also suggested that the double shells often observed in elliptical galaxies, typically ascribed to merging events, can actually be the result of SF in outflows.

The integrated spectra of local elliptical galaxies does not seem to show evidence of a large fraction of stars on radial orbits (Cappellari et al. 2007). However, the study of Cappellari et al. (2007) of 24 local elliptical galaxies results into a radial anisotropy parameter $\beta_r > 0$ (i.e. preference for radial orbits) in ∼2/3 of their sample. The kinematics of individual stars would be needed to properly assess the fraction of the radial orbits.

For galactic bulges and galactic haloes, for which SF within outflows is likely more important, a similar analysis on the distribution of radial and tangential orbits has not yet been performed to our knowledge. Very recently, Belokurov et al. (2019), by using the latest Gaia release, have discovered a large population of Galactic halo stars that are metal rich and have highly radial orbits, which are indeed the properties expected of stars formed in past outflows produced by the Galactic centre. According to their analysis, such metal rich stars on radial orbits formed between 9.5 and 12 Gyr ago. Within this context SF in outflows has also been invoked to explain the population of hypervelocity stars in the Galactic halo (Silk & Mamon 2012; Zubovas et al. 2013c; Wang & Loeb 2018).

5.6.2 BH−galaxy correlations

Clearly, stars formed in AGN-driven outflows can potentially directly produce the $M_{\rm BH} - M_{\rm spheroid}$ relation. In particular, given the correlation between outflow rate and AGN luminosity

(i.e. BH accretion rate), obtained by various observational works (Fluetsch et al. 2018; Fiore et al. 2017; Cicone et al. 2014), and assuming the correlation between outflow rate and SF in the outflow found in section 5.3.3 (figures 5.12 (a) and (b)) holds true for all AGN-driven outflows, it is clear that the combination of the two can result into a correlation between BH mass and stars formed in the outflows.

Quantifying the expected relation is yet not simple, as different authors have found different scaling relations between ionised outflows and AGN luminosity (Fiore et al. 2017; Fluetsch et al. 2019). However, if we consider the compilation by Fiore et al. (2017) (where the correlation between ionised outflow rate and AGN luminosity is slightly superlinear), and assume an AGN radiative efficiency of 0.1, then the relation between AGN luminosity and ionised outflow rate translates into:

$$\log\left(\dot{M}_{\mathrm{outfl-ion}}\right) = 0.85 + 1.29\log\left(\dot{M}_{\mathrm{BH}}\right) \tag{5.8}$$

where both the ionised outflow rate and BH accretion rate are in units of $\mathrm{M}_{\odot}\ \mathrm{yr}^{-1}$, and where we have corrected the outflow rates estimated by Fiore et al. (2017) (specifically by a factor of three) to account for the different assumptions to derive the outflow rate in our work. Combining equation 5.8 with the relation between ionised outflow rate and SF in the outflow, i.e. equation 5.2 gives:

$$\mathrm{SFR}_{\mathrm{outfl-ion}} = 15\left(\dot{M}_{\mathrm{BH}}\right)^{1.56} \tag{5.9}$$

where $\mathrm{SFR}_{\mathrm{outfl-ion}}$ is the SFR inside the outflow, again in units of $\mathrm{M}_{\odot}\ \mathrm{yr}^{-1}$. The relation is superlinear. At low/intermediate accretion rates ($\dot{M}_{\mathrm{BH}} < 5\ \mathrm{M}_{\odot}\ \mathrm{yr}^{-1}$), the relation gives a $\mathrm{SFR}/\dot{M}_{\mathrm{BH}}$ ratio lower than the observed local bulge-to-black hole mass ratio, i.e. $M_{\mathrm{bulge}}/M_{\mathrm{BH}} \sim 200$ (Kormendy & Ho 2013), whilst at high accretion rates ($\dot{M}_{\mathrm{BH}} > 5\ \mathrm{M}_{\odot}$) the implied $\mathrm{SFR}/\dot{M}_{\mathrm{BH}}$ ratio is higher than the local bulge-to-black hole mass ratio. This behaviour in two different regimes can potentially explain the very low $M_{\mathrm{bulge}}/M_{\mathrm{BH}} \sim 30$ observed at high redshift (e.g. Wang et al. 2013; Decarli et al. 2018), which can be reached through the first (low accretion rate) regime, in which the BH forms faster than the stellar population (relative to the local relation), whilst later (in the high accretion rate regime) the implied higher SF in the outflow can enable the galaxy to build up quickly enough stellar mass to reach the local relation. Of course, due to the relations being non-linear and the correlations not yet well established, more work is needed to properly assess exactly whether the BH−galaxy correlations and their evolution can be fully explained through SF in outflows or not.

It is finally interesting to note that Fiore et al. (2017) and Bischetti et al. (2018) have identified a very steep relationship between outflow velocity and AGN luminosity. More

specifically, the outflow velocity (combining molecular and ionised outflows) scales with AGN luminosity, i.e. BH accretion rate, as:

$$L_{\rm AGN} \propto v_{\rm outfl}^{4.6} \tag{5.10}$$

The slope of this correlation is comparable to the slope of the correlation between BH mass and stellar velocity dispersion:

$$M_{\rm BH} \propto \sigma_v^{4.5-5} \tag{5.11}$$

(e.g. Shankar et al. 2016). Therefore, if during the active outflow phases BHs are accreting close to the Eddington limit (hence $L_{\rm AGN} \propto M_{\rm BH}$), as expected by many outflow models (King & Pounds 2015), and stars form inside such AGN-driven outflows, then the stellar velocity dispersion of the resulting spheroidal stellar component would have a velocity dispersion consistent with that of the $M_{\rm BH} - \sigma$ relation, without having to invoke complex negative feedback models.

5.6.3 Extragalactic supernovae, CGM/IGM in-situ enrichment, and halo heating

The population of stars formed in galactic outflows can result in supernovae (SNe) exploding on large orbits, outside the galaxy. As discussed in the previous sections most of the stars formed in the outflow are rapidly decelerated (within a few Myr) and will quickly fall back on to the galaxy, hence a significant fraction of the associated SNe explode close to the galaxy. However, a fraction of stars in the outflow may have larger orbits or even escape the galaxy, as expected in models (Zubovas et al. 2013c) and as observed in chapter 4. The resulting SNe should appear as SNe exploding outside galaxies.

Search for SNe tends to be biased towards SNe exploding inside galaxies, as monitoring and follow-up are typically focused on galaxies. However, hostless or intergalactic SNe have been observed in clusters (Gal-Yam et al. 2003; Graham et al. 2015; Gupta et al. 2016).

It should be reminded that locally SF in outflows is a relatively mild phenomenon; hence, SNe outside galaxies are not expected to be very common. However, at high z, where this phenomenon is expected to be much more prominent, SNe exploding in the CGM or even in the IGM should have been much more common. Such extragalactic SNe would have directly enriched, in-situ, the CGM and IGM, relaxing the need of metals to be expelled from galaxies through winds.

An additional implication is that SNe in the halo would be expanding in a very low-density environment, hence releasing most of the energy to the halo (in contrast to galactic SNe, which

explode in dense environments hence losing most of their energy through radiative losses, Walch & Naab 2015). As a consequence, SNe outside the galaxy can be very effective in heating the halo, hence suppressing the accretion of cold gas on to the galaxy, resulting in a delayed ('preventive') feedback, which may result in quenching SF as a consequence of starvation. Therefore, whilst galactic outflows have a positive initial effect, resulting in new stars born in the outflow, eventually they may actually have a net negative feedback effect on the entire galaxy through halo heating and galaxy starvation.

5.6.4 Reionisation of the Universe

As discussed above, as soon as stars form in the outflow they only respond to gravity, hence they quickly decouple from the outflowing clouds (which are instead also subject to radiation pressure, ram pressure, and other fluidodynamic effects). As a consequence, the massive young, hot stars formed in the outflow will have a very high escape fraction of ionising photons. Therefore, if this phenomenon is taking place in the early Universe, SF in outflows can greatly contribute to the reionisation of the Universe. Indeed, one of the main problems of current theories is that the ionising photons escape fraction typically observed in galaxies is very low (less than 5−10 per cent), making it difficult for models to account for the reionisation of the Universe. The large escape fraction expected for stars formed in outflows may help to tackle this long-standing issue.

5.7 Conclusions

Several models have proposed that massive galactic outflows may form stars inside them, with potentially far reaching implications as stars formed inside outflows would have completely different kinematic properties than those formed in galactic discs. The detection of large amounts of molecular gas, both dense and clumpy, in galactic outflows does indeed support the scenario that galactic outflows should form stars. However, direct evidence of SF inside galactic outflows was so far found only in a single galaxy, leaving unclear whether this is a rare phenomenon or common to galactic outflows.

In order to address this issue, we have analysed integral field spectroscopic data of 2800 local galaxies in the MaNGA DR2. Through a detailed spectral analysis, we have identified a sample of 37 galaxies that shows clear evidence for outflows and whose excitation mechanism can be investigated through the detection of multiple nebular lines, which can be used to trace the excitation of outflowing gas through the BPT diagnostic diagrams.

We have obtained the following results:

(i) About 30 per cent of the outflows have BPT diagnostic diagrams consistent with being SF; that is, the nebular lines in the outflow are primarily excited by the UV radiation of young hot stars. Moreover, we find that half of the outflows show at least some fraction of them (>10 per cent) being excited by star formation. These results highlight that SF is common in a large fraction of galactic outflows.

(ii) The analysis of the ionisation parameter in the star-forming outflows, which is indistinguishable from (or even higher than) normal star-forming regions, confirms that the outflowing gas is not photoionised by the UV radiation coming from the underlying galactic discs, but must be photoionised by *in-situ* (i.e. within the outflow) SF.

(iii) In these local galaxies, the SF inside the outflow generally accounts only for about 5−30 per cent of the total star formation in the galaxy. However, SF inside the outflow can contribute significantly, or even dominate SF in the central few kpc.

(iv) We find that, for those galaxies whose outflow is dominated by SF, the SFR inside the outflows correlates with the ionised outflow rate. If extrapolated to the outflow rates observed in distant massive galaxies and quasars, then the implied SFR inside those outflows can be up to several 100 $M_{\odot}$ yr^{-1} hence potentially contributing significantly to the evolution of galaxies. We suggest that evidence for SF in the outflows of distant galaxies may already be present.

(v) We show that galaxies with SF inside the outflow are primarily distributed along the 'main sequence' (or, in some cases, slightly above it) and are typically isolated, regular disk galaxies.

(vi) We discuss the possible reasons why SF may have remained elusive until recently.We show that even if prominent star formation is present inside the outflow, the presence of an even faint AGN generally dominates the diagnostics. This may have precluded the detection of SF in many outflows, especially in integrated spectra or with poor spatial resolution.We also point out that outflows identified through the [OIII]5008 line may have biased the samples against those dominated by SF. Moreover, the use of the [NII]-BPT diagnostic diagram may also have resulted in improperly classifying the nature of some outflows.

(vii) As soon as stars form inside galactic outflows they react only to gravity, they decouple from the gaseous outflow and move ballistically. Stars formed in galactic outflows may either escape the galaxy, or remain gravitationally bound. As our study has revealed that SF inside galactic outflows is a relatively common phenomenon, we have discussed the implications of such widespread production of stars with kinematic properties completely different from those formed in galactic discs. Specifically:

(a) SF inside outflows can potentially contribute significantly to the formation of the spheroidal component of galaxies (bulge, halo, and possibly also part of elliptical galaxies). Within this context, stars formed inside galactic outflows can explain the population of stars on radial orbits recently discovered in our Galaxy, as well as the population of hypervelocity stars. The same phenomenon can explain she double-shells observed in elliptical galaxies.

(b) This phenomenon can contribute to establish the correlations between BHs and their host galaxies, and in particular the $M_{\rm BH} - \sigma$ and the $M_{\rm BH} - M_{\rm spheroid}$ relations.

(c) Stars formed in galactic outflows and escaping the galaxy would produce SNe exploding outside galaxies, hence accounting for the population of host less SNe observed in galaxy cluster. SNe exploding on large orbits, or outside galaxies, can contribute to the in-situ enrichment of the CGM and of the IGM. Such SNe can also result in a very effective heating of the galactic halo, hence preventing cold gas accretion on to the galaxy and therefore contributing to the delayed quenching of SF in the galaxy as a consequence of starvation.

(d) Young stars formed in the outflow quickly decouple from their molecular cloud and therefore are expected to have large escape fraction of ionising photons. As a consequence, if star formation in outflows is common to primeval galaxies, this phenomenon can contribute significantly to the reionisation of the Universe.

PID	IFU	RA [deg]	DEC [deg]	z	log (M_{star}) [$M_\odot$]
7443	12703	229.52558	42.74584	0.04027	10.82[a]
8241	6102	126.05963	17.33195	0.03725	10.45
8244	3702	131.81500	51.24580	0.02754	10.01
8252	1902	146.09183	47.45985	0.02589	9.35
8256	12704	166.12940	42.62455	0.12610	11.16[a]
8318	6102	197.23931	45.90544	0.12907	11.34
8329	3701	213.43219	43.66248	0.08934	–[a]
8341	12704	189.21325	45.65117	0.03034	10.62
8439	6102	142.77816	49.07974	0.03393	10.72
8482	12704	243.58182	50.46561	0.06025	11.39
8550	3704	248.42638	39.18511	0.02984	10.44
8550	9102	247.20905	39.83508	0.03585	10.67
8588	6101	248.45675	39.26320	0.03176	10.74
8612	12704	254.56457	39.39146	0.03431	10.76
8715	3702	119.92067	50.83997	0.05436	10.33
8931	9101	193.07383	27.08555	0.02098	10.25
8946	3701	168.95772	46.31956	0.05328	10.67
8948	12704	167.30601	49.51943	0.07242	11.02
8952	3703	205.44092	27.10634	0.02878	10.33
8979	6102	241.82338	41.40360	0.03463	10.76
9026	9101	249.31841	44.41822	0.03141	10.82
9049	1901	247.56097	26.20647	0.13145	11.38
10001	6102	132.65399	57.35966	0.02610	10.67
8250	12704	138.98137	44.33276	0.03978	10.95[a]
8263	3701	184.76431	46.10679	0.03856	10.09
8257	12701	165.49581	45.22802	0.01999	10.38
8464	3702	187.0635	44.45313	0.02289	–[a]
7992	6104	255.27948	64.67687	0.02713	10.21
8078	6104	42.739429	0.369408	0.04421	10.05
8082	12701	48.896456	-1.01628	0.02680	10.44
8249	3703	139.72046	45.72778	0.02643	9.79
8250	3703	139.73996	43.50057	0.04005	9.62
8262	9102	184.55356	44.17324	0.02452	10.29
8315	12705	235.92048	39.54035	0.06347	10.60
8550	12703	247.67443	40.52938	0.02981	10.55
8726	12701	115.71703	22.11273	0.02864	10.72
8931	12702	192.76494	27.36996	0.02772	10.07

Table 5.1: List of MaNGA galaxies whose outflow can be characterised through the BPT diagrams. The columns provide the following information: 1) Plate ID; 2) IFU ID; 3) Righ Ascension; 4) Declination; 5) Stella Mass taken from the MPA-JHU catalog (Brinchmann et al. 2004) (lack of entry means that the stellar mass is not available in the MPA-JHU catalogue).
[a] Interacting systems.

PID	IFU	SFR(disk) ($M_\odot$ yr^{-1})	SFR(outflow) ($M_\odot$ yr^{-1})	$\dot{M}_{outf}$(Hα) ($M_\odot$ yr^{-1})	$\dot{M}_{outf}$([OIII]) ($M_\odot$ yr^{-1})	BPT Class. disk [NII]	[SII]	[OI]	BPT Class. Outf [NII]	[SII]	[OI]
7443	12703	36.38	–	2.608	0.057	Comp	SF	LIER	AGN	AGN	AGN
8241	6102	5.21	–	0.37	0.036	Comp	SF	SF	AGN	AGN	AGN
8244	3702	1.39	–	0.09	0.004	SF	SF	SF	AGN	AGN	AGN
8252	1902	0.07	–	0.01	0.001	SF	SF	LIER	AGN	AGN	AGN
8256	12704	2.45	–	0.59	0.038	AGN	LIER	AGN	AGN	AGN	AGN
8318	6102	9.13	–	1.99	0.194	Comp	SF	AGN	AGN	LIER	LIER
8329	3701	0.85	–	0.15	0.015	Comp	SF	LIER	AGN	LIER	LIER
8341	12704	2.45	–	0.21	0.045	SF	SF	SF	AGN	AGN	AGN
8439	6102	4.98	–	0.61	0.023	Comp	SF	SF	AGN	AGN	AGN
8482	12704	1.22	–	<0.1	0.014	AGN	LIER	AGN	AGN	LIER	LIER
8550	3704	0.19	–	0.34	0.012	AGN	LIER	SF	AGN	LIER	LIER
8550	9102	4.90	–	0.07	0.004	SF	SF	SF	AGN	AGN	AGN
8588	6101	3.82	–	0.24	0.013	Comp	SF	SF	AGN	LIER	AGN
8612	12704	1.76	–	0.16	0.015	AGN	AGN	AGN	AGN	AGN	AGN
8715	3702	2.70	–	16.90	0.887	AGN	AGN	AGN	AGN	AGN	AGN
8931	9101	0.36	–	0.15	0.001	Comp	SF	LIER	AGN	AGN	AGN
8946	3701	0.51	–	0.06	0.005	Comp	SF	SF	AGN	AGN	AGN
8948	12704	4.43	–	0.17	0.013	Comp	SF	SF	AGN	AGN	AGN
8952	3703	0.24	–	0.02	0.002	Comp	SF	SF	SF	LIER	LIER
8979	6102	1.37	–	0.05	0.004	SF	SF	SF	AGN	LIER	AGN
9026	9101	1.23	–	0.20	0.008	Comp	SF	SF	AGN	AGN	AGN
9049	1901	15.24	–	0.49	0.033	Comp	SF	SF	AGN	AGN	AGN
10001	6102	1.17	–	0.24	0.004	Comp	SF	SF	AGN	AGN	AGN
8250	12704	16.57	4.30	0.33	0.009	Comp	SF	SF	AGN	SF	AGN
8263	3701	0.61	0.31	0.13	0.003	Comp	SF	SF	AGN	SF	LIER
8257	12701	5.824	0.64	0.22	0.016	Comp	SF	SF	AGN	SF	SF
8464	3702	2.20	0.17	0.12	0.010	Comp	SF	SF	AGN	SF	SF
7992	6104	2.86	0.29	0.09	0.003	Comp	SF	SF	Comp	SF	SF
8082	12701	1.10	0.09	0.06	0.004	SF	SF	SF	Comp	SF	SF
8262	9102	5.00	0.93	0.36	0.051	SF	SF	SF	Comp	SF	SF
8315	12705	5.74	1.04	0.26	0.021	SF	SF	SF	Comp	SF	SF
8550	12703	1.11	0.09	0.04	0.004	SF	SF	SF	Comp	SF	SF
8726	12701	0.93	0.23	0.11	0.009	Comp	SF	SF	Comp	SF	SF
8931	12702	1.72	0.10	0.05	0.003	SF	SF	SF	Comp	SF	SF
8078	6104	1.32	0.88	0.24	0.027	Comp	SF	SF	SF	SF	SF
8249	3703	0.65	0.03	0.01	0.002	SF	SF	SF	SF	SF	SF
8250	3703	1.42	0.29	0.07	0.016	SF	SF	SF	SF	SF	SF

Table 5.2: Properties of the SF rates, outflow rate and BPT classification of the discs and of the outflows. The columns provide the following information: 1) Plate ID; 2) IFU ID; 3) SF Rate in the disk; 4) SF Rate in the outflow (only for those galaxies that are dominated by SF in the outflow according to the [SII]-BPT diagram); 5) Ionised outflow rate determined through the broad Hα emission; 6) Ionised outflow rate determined through the broad [OIII] emission; 7–9) BPT classification of the nebular emission of the galaxy disk; 10–12) BPT classification of the nebular emission of the outflow.

Chapter 6

The properties of ionised outflows and their host galaxies

This chapter is adapted from 'The properties of ionised outflows and their host galaxies', R. Gallagher et al. in prep.

The availability of gas reservoirs in the inner kiloparsecs of galaxies strongly drives the evolutionary paths of galaxies throughout cosmic time. These reservoirs of gas serve as fuel for accretion onto SMBH in galaxies hosting AGN or for the perpetuation of SF, yet are affected by negative feedback due to the energy and momentum injected by AGN supernovae explosions and stellar winds. As discussed in chapter 1 these feedback mechanisms can directly impact the ISM, disrupting, heating and mixing the gas and, potentially, expelling it from the galaxy as outflows, and their effects are required to explain a number of present day scaling relations between BH mass and host galaxy properties, such as the $M_{BH} - \sigma_*$ relationship, and prevent the overgrowth of galaxies (e.g. Silk & Rees, 1998; Fabian 2012; King & Pounds, 2015). Hence, a deep understanding of the incidence and properties outflows and their feedback effects is required to truly determine their roles in galaxy evolution.

Both AGN- and SF-driven outflows are found to ubiquitously present in local (Veilleux et al. 2005; Chen et al. 2010) and high redshift galaxies (Rubin et al. 2014). Typically, galactic outflows are induced as a 'negative feedback' mechanism is galaxies, as outflows hydrodynamically coupled to the ISM can induce feedback effects on their host to slow and potentially halt ongoing SF through the removal and heating of the SF gas reservoirs in the ISM,

subsequently preventing the stellar mass growth of galaxies (Veilleux et al., 2005; Heckman & Borthakur 2016; Cresci & Maiolino 2018).

Typically, outflows are traced through gas moving at high velocities in comparison to their host's systemic gas component. In the ionised phase, outflows are typically traced through either the Hα or [OIII]λ5008 line emission of the low density (n_e $\sim 10^2 - 10^4$ cm^{-3}, e.g. Vaona et al. 2012) ionised gas entrained within outflows. This ionised component of the outflow may account for anywhere from only a few percent to up to 50% of the total outflow mass for AGN or SF galaxies, respectively (Fluetsch et al. 2019). However, even in those cases where the ionised component constitutes only a minor part of the outflow, the optical nebular lines provide some of the strongest diagnostic tools to trace and map outflows in large samples of galaxies.

Whilst in chapters 4 and 5 we have focused heavily upon the 'positive feedback' mechanisms of outflows, in this chapter we instead focus upon the 'negative feedback' effects of outflows and, more generally, the physical properties of outflows and their occurrence in local galaxies. We present our investigations into the propensity and properties of outflows, as well as their effects on their hosts, for local Universe galaxies by exploiting the IFU data of ~4600 galaxies from the MaNGA survey. We investigate the properties of the outflow host galaxies as a whole, as well as the powerful outflow driving sources they host, through emission line diagnostics. We also investigate the properties of the outflowing gas. One of the key differences between this chapter and chapter 5 is the we no longer require the detection of *all* of the nebular lines necessary for BPT diagnostics at sufficient SNR within the outflowing gas. Here, instead, we consider all galaxies with sufficient outflow signatures in at least one of Hα and [OIII]λ5008, resulting in a larger, less biased sample that provides a stronger statistical framework from which to work.

6.1 Data Reduction

For this chapter, we make use of the full MaNGA MPL-7 dataset (equivalent to the data in SDSS DR15, Aguado et al. 2018) consisting of 4621 unique galaxies. The main draw to change to this new dataset was the introduction of the a new extension to the datacubes containing the wavelength dependent line spread function (LSF) at each spaxel. This improvement to the LSF estimate allows for better recovery of both stellar and gas kinematics.

To analyse the data, we make full use of the fitting procedures offered by SCRAPPy, which utilises the `VORBIN` algorithm (Cappellari & Copin, 2003) to spatially bin data and increase SNR (to which we provide the custom SNR function shown in section 3.3.2) and a combination of `pPXF` (Cappelari, 2017) and custom algorithms to recover the stellar continuum and kinematics, and emission line properties. We run SCRAPPy fully, through to the smoothed spaxel input

fitting module. We also make use of the stellar population synthesis module of SCRAPPy to recover the stellar masses (only in binned spectra, as individual spaxels are likely to contain insufficient SNR for accurate recovery) as well its ability to recover the D4000 and D_n4000 stellar indices. For the binning, we require a minimum bin SNR of 10 per spectral pixel on the continuum within the SDSS r-band, with an initial SNR cut of 3 on the continuum chosen to minimise contamination from low signal regions.

To make use of the stellar population synthesis module, we provide SCRAPPy with a template library consisting of SSPs from the MILES library (Falcón-Barroso et al. 2011). The templates from this library represent single-age, single-metallicity stellar population models, computed from the MILES empirical spectra. They are given in units of $L_\lambda/L_\odot M_\odot^{-1}\text{Å}^{-1}$, and the total mass of each SSP is $1M_\odot$. Thus, for a fit representing the underlying population of a spectrum, the contained stellar mass can be recovered from the weighs of the contributing SSPs.

Taken directly from the MILES website, we select a subset of the available SSPs covering 24 ages, from 79 Myr to 15.8Gyr in logarithmically spaced bins, and 6 metallicites, covering [Z/H] = -1.71 to [Z/H] = +0.22. These templates were based upon the Padova isochrones (Girardi et al. 2000), and computed with unimodal IMF with a slope of 1.3, equivalent to the Salpeter IMF. These ranges of values and input ingredients were chosen to ensure all templates fell within the "SAFE" ranges for the MILES models (see the MILES website http://miles.iac.es/ or Vazdekis et al. (2010) for more details). This template library is used throughout all fitting modules.

For the kinematics modules, where our main concerns are accurate recovery of the stellar and emission line kinematics, we make use of both an 8^{th} order additive and a 4^{th} order multiplicative polynomial to correct the continuum shape for flux calibrations issues. However, for the population synthesis, we only employ multiplicative polynomials, of 8^{th} order, to account for flux calibration errors, as recommended by Cappellari & Emsellem (2004). This is because additive polynomials may alter line strengths within the fit and bias the inferred ages and metallicities and, though we do not investigate these properties, SSPs of different properties will contribute more or less light to the total recovered continuum, and thus provide more or less stellar mass. As such, we wish to avoid bias where possible. For the population fits we require SCRAPPy to apply a 2^{nd} order linear regularisation constraints to the weights of the fits. Inspired by previous works that have explored recovery of stellar population parameters in this fashion and on large scales (McDermid et al. 2015; Guérou et al. 2015; Guérou et al. 2016) we determine the level of the applied regularisation through fitting the central extracted spectrum of each galaxy. The degree of regularisation is determined iteratively by SCRAPPy, then held constant for all other spectra of the galaxy. This process ensures we end up with the

smoothest distribution of ages and metallicities, whilst remaining statistically consistent with the data.

For the emission lines, we used all of the default settings for SCRAPPy: specifically, we require a minimum SNR of 3 and Equivalent Width (EW) of 3Å in at least one of the Hα or [OIII]λ5008 broad emission line components, with the broad component improving the reduced chi-squared value of narrow component fit by a minimum of 10%.

6.2 Spatially resolved diagnostics of outflow host galaxies

We begin our investigations into the properties of our outflow host galaxies by considering the nebular line emission of the systemic (narrow) component of our fits. A number of major physical mechanisms can be invoked to explain the presence of ionised outflows, such as supernovae, stellar winds, shocks and AGN activity, all of which will affect the ISM of the galaxy. Furthermore, we wish to determine regions in our galaxies exhibiting ongoing SF, quiescence and AGN presence. As previously mentioned in chapter 1, the most common approach used in spectroscopic surveys to distinguish different ionisation sources and excitation mechanisms are diagnostic diagrams that utilise ratios of the nebular emission lines. Here, we make use of the BPT diagnostic diagrams (Baldwin, Phillips & Terlevich 1981), supported by the WHAN diagnostic diagram (Cid Fernandes et al. 2010, 2011). Of note, however, is that we only consider the systemic gas component, not the outflow gas component. Analysis of the latter is achievable only for the small number of our outflows with sufficient SNR in all of the required nebular line emission of the entrained outflow gas, whereas the former can be analysed in all of our outflow host galaxies.

Whilst we use the spatially resolved maps of the diagnostic diagrams to determine regions of SF and quiescence, we also wish to use the traditional diagrams to determine the overall primary ionising source in each galaxy by considering the median location on the diagrams. To achieve this, we consider each of the diagnostics in three defined regions for the systemic component of the host galaxy, traced by the narrow component in our fits. We firstly consider the locations in these diagrams of the host galaxy as a whole, enabling us to investigate the outflow detection rate as a function of host galaxy type. We define this selection as our "All Spaxel" (AS) classification. However, the median locations can be biased by contributions of spaxels from extreme regions, and on a galactic scale only tell us the main source of ionisation across the entire galaxy. As we also wish to understand the feedback effects of the outflows and their driving sources on the galaxies, we further consider the narrow component of only those spaxels that trace the outflow in each galaxy, which we denote as our "Outflow Total" (OT) spaxels. Finally, under the assumption that the outflows are driven centrally from a single

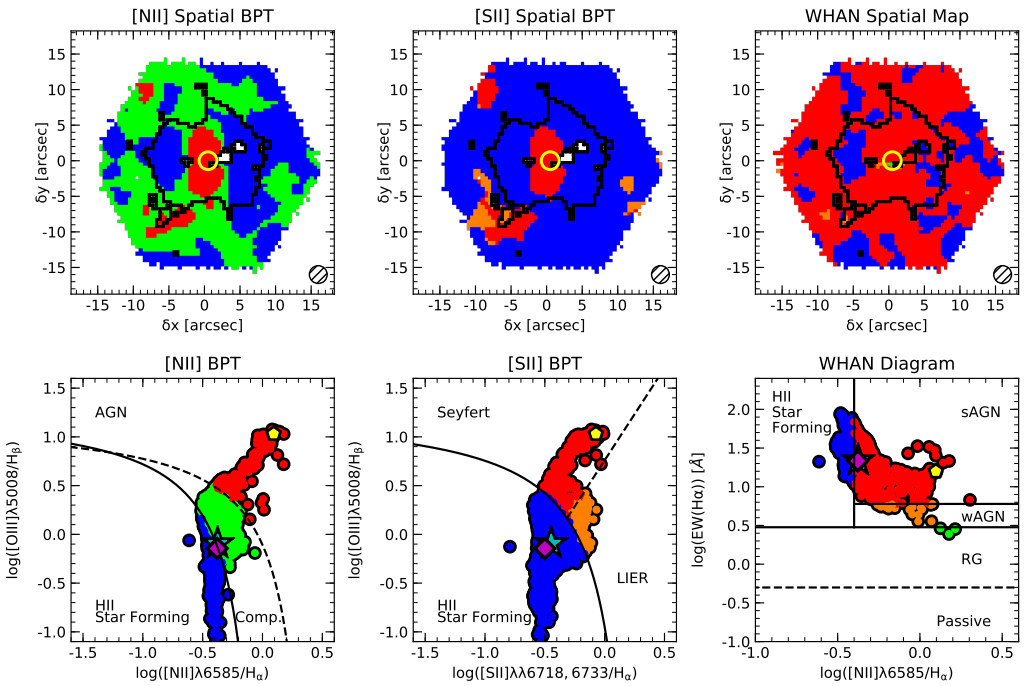


Figure 6.1: Upper panels: Spatially resolved BPT and WHAN classifications of the narrow component gas. Shown as the black bounding region are the spaxels that have an outflow component detected in their emission lines. The yellow circle indicates the spaxels contained within an aperture of the size of the MaNGA PSF, centred on the outflow luminosity weighted centroid spaxel, and assumed to trace the outflow driving source. The hashed circle in the bottom right corners shows the size of the MaNGA PSF. Lower panels: Distributions of the narrow component nebular lines on the BPT and WHAN diagrams. The median classifications of all of the spaxels, all spaxels with outflow component detections (black bounded box on the diagrams above) and the luminosity centred PSF spaxels (yellow circle in the diagrams above) are shown as the magenta diamond, blue star and yellow pentagon, respectively.

location, we consider those spaxels that fall within an aperture of the size of the MaNGA PSF, centred on the Hα and [OIII] median outflow (broad) component luminosity weighted centroid spaxel. This classification of spaxels, henceforth referred to as "Outflow Central" (OC) spaxels, are selected to attempt to determine the driving source of the outflows. Figure 6.1 shows the spatially resolved maps and standard diagnostic diagrams of all three diagnostic tools for a sample galaxy, with the OT and OC spaxel selections shown on the spatially resolved maps (upper panels) as the black, bounding region and the yellow circle, respectively. The effect of considering the three defined spaxel classifications is highlighted in the standard diagnostic diagrams, shown in the bottom panels, with the median locations of our AS, OT and OC spaxels shown as a magenta diamond, a blue star and a yellow pentagon, respectively. As can be seen,

the AS and OT median classifications fall in a similar location, yet the OC median is located strongly within the AGN loci of both diagrams. Similarly, in the WHAN diagram, whilst the AS and OT median locations are classified as sAGN, they are close to the SF demarcation line. Conversely, the OC median is located approximately 0.5 dex away from the SF demarcation. This suggests that the final classification is certainly needed to truly disentangle the driving source of the outflow from other ionisation sources in the galaxy.

The strong disagreement between the BPT diagrams and the WHAN diagram over a majority of the galaxy suggests that the WHAN diagram may not be as useful a diagnostic tool on spatially resolved scales. The SF demarcation line of the WHAN diagram is a translation of the demarcation line from Stasińska et al. (2006), which serves to separate pure-SF regions from those that may contain ongoing SF but are contaminated by AGN. Furthermore, similarly to the [NII]-BPT diagram, the WHAN diagram is affected by both the nitrogen abundance ratio and ionisation parameter, which can mean regions with ongoing SF my be shifted into the sAGN locus. As such, although we consider the properties of the WHAN diagnostics of our outflow host galaxies, we note that the conclusions drawn are likely less reliable than those of the BPT diagrams.

6.2.1 Spatially resolved BPT diagrams

We start our investigations with the spatially resolved BPT diagnostic diagrams. In figure 6.2 we show the resulting median locations on the [NII]- and [SII]-BPT diagrams for the AS, OT and OC spaxels for all galaxies hosting outflows. Also shown in the background of each plot, as grey hexbins coloured by local density, are the results from the SDSS DR7 dataset, with specific cuts made to the dataset for consistency with the MaNGA galaxy selection criteria.

We begin by comparing the population levels of our outflow hosts in each of our three subregions in comparison to the local density of the SDSS data, summarised in table 6.1. Starting with the AS classification, from both the [NII] and [SII] diagnostics we see that our outflow hosts seem representative of galaxies within the local Universe. Considering our method for locating outflows is unbiased (aside from our initial selection criteria of requiring a nuclear outflow), this suggests that outflows are prevalent in all types of galaxy in the local Universe. However, we do note here that there are a number of systematic differences between our median BPT classifications and the classifications from the SDSS data. Whereas our data is spatially resolved, and therefore a better indicator of the general properties of the galaxy, the DR7 observations were made using a single 3 arcsecond fibre trained onto the centre of each of the targeted galaxies. Whilst a full discussion of the differences between the two methodologies is outside of the scope of this thesis, the differences have been explored in previous works (i.e. Belfiore et al. 2016), with the general consensus that, although the spatially resolved

Data Set	Count Type	[NII] BPT Classification			[SII] BPT Classification			No. Galaxies
		SF	Comp.	AGN	SF	LIER	Sy	Total
SDSS DR7	Number	254752	71965	42847	315201	35986	18377	369564
	% of Total	68.9%	19.5%	11.6%	85.3%	9.7%	5.0%	
All Spaxels	Number	107	48	13	151	9	8	168
	% of Total	63.7%	28.6%	7.7%	89.9%	5.4%	4.8%	
Outflow Total Spaxels	Number	105	40	23	142	14	12	168
	% of Total	62.5%	23.8%	13.7%	84.5%	8.3%	7.2%	
Outflow Centre Spaxels	Number	86	46	36	131	18	19	168
	% of Total	51.2%	27.4%	21.4%	78.0%	10.7%	11.3%	

Table 6.1: Classification population counts for median BPT diagnostic diagram classifications. Shown are both the [NII] and [SII] median BPT diagnostic populations for the SDSS DR7 galaxies, with SNR>3 in all lines, as well as the three spaxel selection criteria discussed in section 6.2. Specifically, shown are the median classification populations when considering all disk spaxels, only those spaxels with outflow detections, and only those spaxels within 1 PSF of the luminosity centroid of the outflow.

information provides a better view of the galaxy processes individually, the general shape and populations explored by the SDSS data on the BPT diagrams sufficiently represent the true distributions of the local Universe.

Whilst this view provides insight into the general classifications of outflow host galaxies, it provides less information about the driving mechanism of the outflow itself. Furthermore, as we are exploring the median locations of all spaxels on the BPT diagrams, the classifications can be biased against obscurred regions in the galaxy (especially obscuration of the nucleus) due to inclination. The impact of such effects on classifying galaxies by their spatially resolved diagrams was explored in Wylezalek et al. 2018, wherein they attempted to explore the AGN population within the MaNGA MPL-5 dataset. To overcome these issues, they introduced an AGN selection scheme based around the fractional number of spaxels in the [NII] and [SII]-BPT diagrams that fell into the AGN loci, coupled with cuts to the EW(Hα) and Hα surface brightness. Although we do not employ the same selection criteria as in Wylezalek et. al (2018), our use of the WHAN diagram (discussed in section 6.2.2, and more specifically the use of EW(Hα) to distinguish areas of true AGN ionisation from those of "fake" AGN ionisation from pAGB star and DIG, was inspired by their work.

As the outflow driving source will also be a constituent part of the host galaxy, we again consider the narrow component of the nebular emission lines. We choose to consider only the narrow component to determine our driving source, rather than the broad component tracing the outflowing gas itself, due to potential ionisation from in situ processes such as SF (Maiolino et al. 2017; Gallagher et. al 2019; Venturi et al. 2019) and shocks.

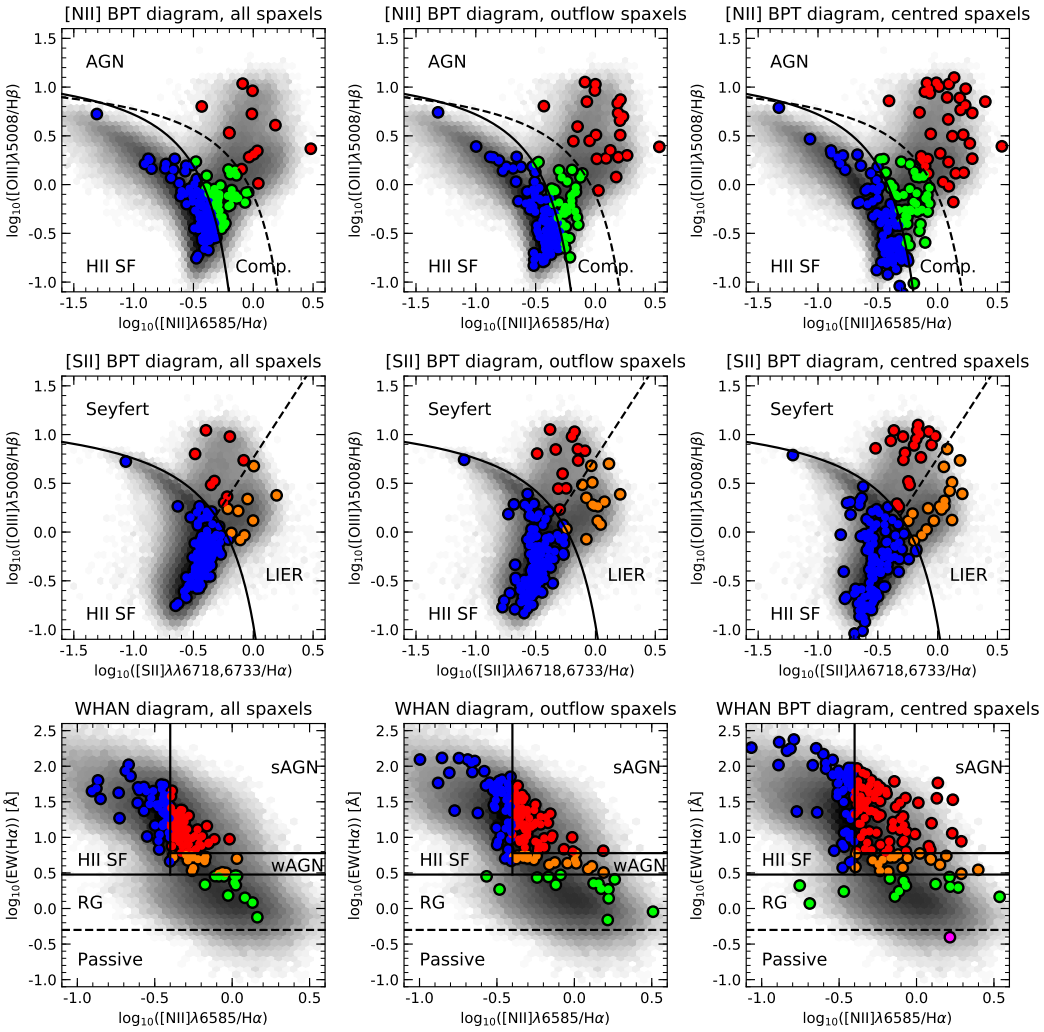


Figure 6.2: Top panels: The median [NII]-BPT classifications for our AS, OC and OT spaxel classifications (left, middle and right panels, respectively), as defined in the text. Middle panels: The median [SII]-BPT classifications for our AS, OC and OT spaxel classifications. Bottom panels: The median WHAN classifications for our AS, OC and OT spaxel classifications. As can be seen, more so in the BPT diagrams than the WHAN diagram, as we consider the spaxels closer to the outflow centre, the number of AGN we trace increases, indicating the important role of AGN in driving outflows.

Firstly, we consider the median BPT locations of the narrow component of our OT spaxels in each galaxy. From table 6.1 it can be seen that for both the [NII] and [SII] populations the number of AGN-like ionising sources increases when compared to the median for the whole galaxy. This increase could be explained by either strong AGN presence or shocks propagating through the ISM. Secondly, we consider the median BPT locations of the narrow component in our OC spaxels. Similarly to considering the OT spaxels, these centroidal spaxels suggest, unsurprisingly, that many of our outflows are AGN-driven. In comparison to the SDSS data, for both the [NII] and [SII]-BPT populations the number of AGN/Sy excitation sources is double that of the general galaxy population. This suggests that outflows are significantly more likely an occurrence in AGN host galaxies. However, there is still a large occurrence of outflows in galaxies dominated by SF in these central spaxels, suggesting SF-driven outflows are more prevalent in the local Universe.

6.2.2 Spatially resolved WHAN diagrams

Additionally, we investigate the spatially resolved WHAN diagrams (Cid Fernandes et al. 2011) for the narrow components of our fits. As in chapter 5 we investigated SF occurrence within the outflows themselves (*in-situ* SF), this method was not possible as measurements of equivalent widths would have relied upon determining and disentangling the stellar continuum of the stars formed in situ from that of the host galaxy. As discussed in chapter 4, the continuum of the putative stars formed inside the outflow is much less luminous than the host galaxy continuum, making it difficult to disentagle. However, as here we consider the properties of the host galaxy alone, we do not need to worry about any external contribution to the continuum, as it will be minimal if present at all.

Figure 6.2 shows the median locations of our three spaxel location choices on the WHAN diagram. Again, the grey hexbins in the background of each plot are the results of the SDSS DR7 dataset. An important advantage of the WHAN diagram is that we only require a minimum SNR of 3 on Hα and [NII]λ6585, thus we have more data points for the WHAN diagram and relative population analysis.

We start by considering our AS spaxel classifications. The population levels of these spaxels are summarised in table 6.2, in comparison to the SDSS DR7 populations. We see a large excess of galaxies classified as strong AGN (sAGN) candidates in comparison to the parent SDSS sample, whilst the SF and weak AGN (wAGN) samples remain consistent. We also see a massive reduction in retired galaxy (RG) and passive galaxy populations. However, due to our initial outflow detection requirements of an SNR ≥ 3 and one of EW(Hα) or EW([OIII]) ≥ 3Å, as discussed in section 6.1, the reduction in these populations is to be expected from selection effects. These results indicate that the WHAN diagnostic is able to distinguish a larger fraction

Data Set	Count Type	WHAN Classification					No. Galaxies
		SF	sAGN	wAGN	RG	Passive	Total
SDSS DR7	Number	186405	104628	45451	120825	4331	461640
	% of Total	40.4%	22.7%	9.8%	26.2%	0.9%	
All Spaxels	Number	61	75	19	13	0	168
	% of Total	36.3%	44.6%	11.3%	7.7%	0.0%	
Outflow Spaxels	Number	65	70	19	14	0	168
	% of Total	38.7%	41.7%	11.3%	8.3%	0.0%	
Outflow Centre Spaxels	Number	55	80	19	13	1	168
	% of Total	32.7%	47.6%	11.3%	7.7%	0.6%	

Table 6.2: Classification population counts for median WHAN diagnostic diagram classification. Shown are the median WHAN diagnostic populations for the SDSS DR7 galaxies, with SNR>3 in all lines, as well as the three spaxel selection criteria discussed in section 6.2.1. Specifically, shown are the median classification populations when considering all disk spaxels, only those spaxels with outflow detections, and only those spaxels within 1 PSF of the luminosity centroid of the outflow.

of outflows associated with AGN than the BPT diagnostics, even at the global level, though this could be in part due to the way in which the loci of the WHAN diagrams are defined, as discussed in section 1.4.2.

We next consider our OT and OC spaxel classifications. We can see from table 6.2 that the populations of the galaxies when considering these locations in comparison to the AS spaxels seems to vary very little, unlike for the BPT populations. This suggests that although the diagnostic is useful for differentiating between sAGN, wAGN and other mechanisms that give rise to LIER like ionisation, it may not be as useful as tool to consider the spatially resolved mechanisms at play within a galaxy on a spatially resolved scale, or that the WHAN classification is incredibly effective, and thus saturates, in identifying AGN at the global scale.

For the WHAN diagnostic the number of sAGN traced is almost double that traced by the SDSS DR7. This suggests that a many of our outflows are AGN-driven, as we would expect. However, a large number of the outflow driving sources ($\sim$30$-$40%) remain classified as SF, complementing the results from the BPT diagrams that SF-driven outflows may be prevalent in the local Universe.

6.3 Distribution of galaxies on the SFR-M_* plane

As discussed in chapter 1, many properties of galaxies can be inferred from their locations on the SFR vs stellar mass (M_*) diagram. For us to investigate our galaxies locations on the diagram, we must begin by recovering the global SFR and M_*. As discussed in section 6.1, we

have recovered the stellar mass surface density Σ_* in each spaxel, and thus can simply integrate over all spaxels to recover the global stellar mass. Recovery of the SFR, however, requires a combination of techniques.

6.3.1 Recovery of the global SFR

Whilst SFRs can be trivially recovered from SF regions with strong line emission from their Hα luminosity, SFRs for regions that are either apparently quiescent (with SF too low to be measured) or show AGN presence (and thus AGN contamination) cannot be recovered via emission line fluxes alone. Furthermore, although subtraction of AGN contribution is possible, the procedures used carry uncertainties. Brinchmann et al. (2004) encountered this issue with the SDSS data, and to overcome the problem searched for a tracer that correlated to SFR in SF galaxies, yet was universally measurable. They opted to use the relationship between the specific SFR (sSFR, SFR per unit mass) and D4000 to determine SFRs for apparently quiescent galaxies and AGN. More recently Bluck et al. (2019) have investigated this relationship on a spatially resolved scale. Both studies agree that D4000 is a good alternative tracer for SFR as it:

(i) correlates well with Σ_{SFR} in star-forming regions
(ii) is effective at identifying quenched regions
(iii) can be measured reliably in all spaxels

Bluck et al. (2019) found that the relationship holds on the spaxel level and holds for both D4000 and D_n4000. Of note, however, is that D4000 (and D_n4000) provide information on the SFR averaged over a longer timescale relative to that traced by Hα.

As SCRAPPy recovers both these stellar indices and stellar mass surface densities, we can combine use the relationships from Bluck et al. (2019) to recover SFRs in apparently quiescent or AGN regions, and in turn combine these SFRs with those determined by Hα luminosities in SF regions, providing us a global SFR for each galaxy. To use this combinational technique, however, first requires we determine regions of pure SF, and regions with either none or so little SFR that we cannot trace it, or AGN contamination. The latter of these two types of regions will henceforth be referred to as "quiescent" regions, though we do note that, as stated, these regions may have actually possess ongoing or recent SF.

To achieve this, we return to our spatially resolved BPT maps. We consider both the [NII] and [SII]-BPT maps in conjunction, in turn allowing us to adopt a scheme for separating pure SF spaxels from those potentially subject to AGN contamination. For a spaxel to be defined as pure SF, we require that the spaxel be classified as SF by the [SII]-BPT diagnostic, and either SF or Composite by the [NII]-BPT diagnostic. Any spaxel which does not meet these criteria

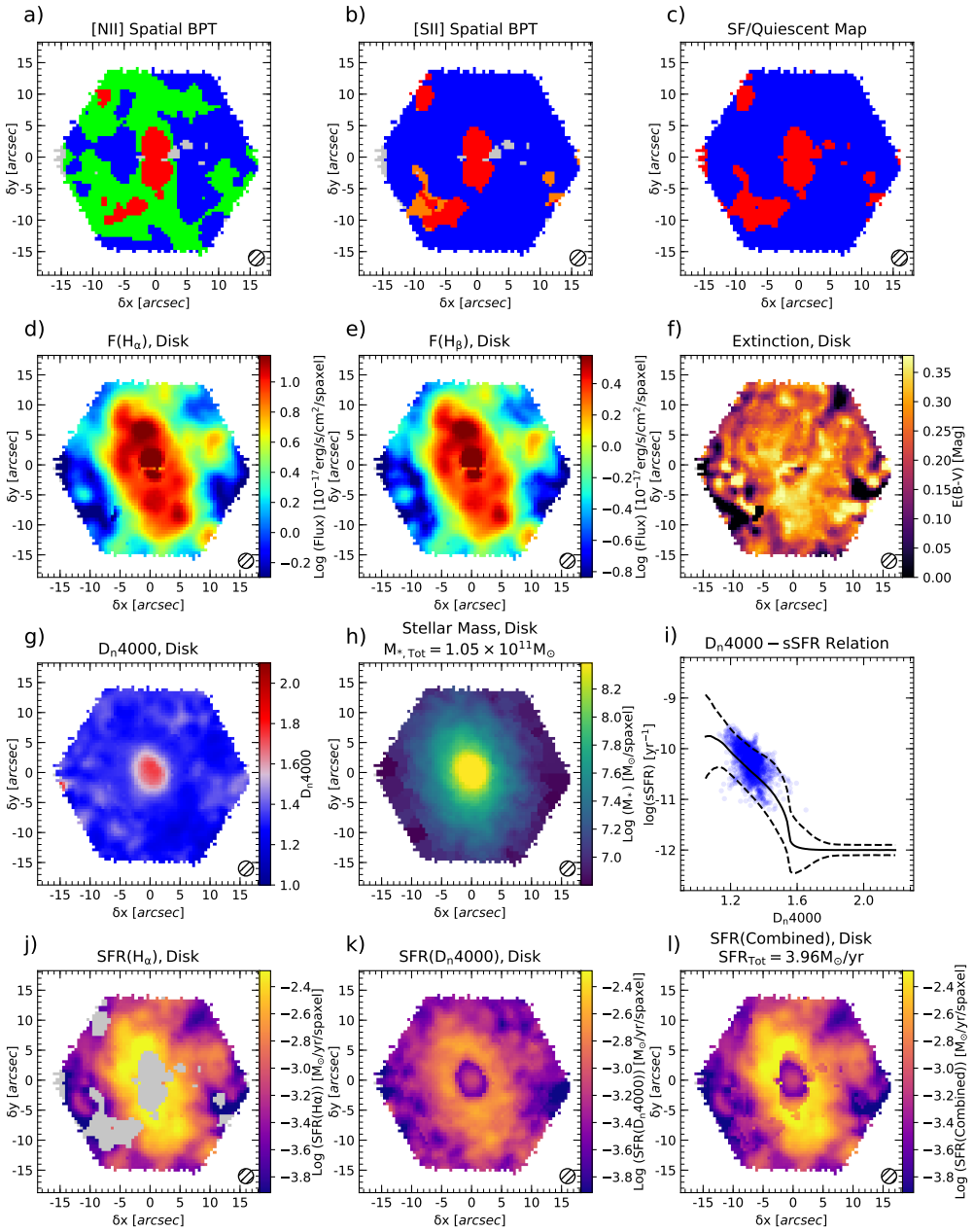


Figure 6.3: Maps used in the determination of a galaxies total global SFR and stellar mass. (a) and (b): spatially resolved [NII] and [SII] maps respectively, colour scheme remains consistent with figure 6.1. (c): map of SF or quiescent spaxels (blue or red respectively), as determined via the [NII] and [SII]-BPT maps. (d) and (e): Flux maps for Hα and Hβ, using to determine reddening. (f): Extinction map derived from reddening from Hα/Hβ ratio, assuming case B recombination. (g): D$_n$4000 index map. (h): Stellar mass map, given as spaxel surface density. Total galaxy mass is stated above the map. (i): Relationship from Bluck et al. (2019) to determine sSFR from D$_n$4000, shown as the solid black line. The dashed lines represent the 1σ confidence bounds. Shown as blue scatter points are the sSFRs determined for each spaxel where SFR can be computed from $L(H\alpha)$, showing the vast majority of points falling within $\pm 1\sigma$. (j) SFR derived from $L(H\alpha)$ using the Kennicutt et al. (1998) law. (k): SFR computed from combining sSFR calculated via D$_n$4000 and stellar mass. (l): Final total SFR map for the galaxy. Total global SFR is shown above the map.

is thus defined as "quiescent" (or AGN). Panels (a) and (b) in figure 6.3 show both the [NII] and [SII]-BPT maps for our showcase galaxy, alongside panel (c) which shows the spatially resolved map of our selection scheme.

For our SF spaxels, we start by correcting our Hα fluxes for extinction using the Calzetti et al. (2000) attenuation curve, assuming case B recombination (Osterbrock & Ferland (2006)), to infer reddening. Panels (d)-(f) of figure 6.3 show the flux maps of Hα and Hβ, and the computed extinction map for the same sample galaxy as in figure 6.3. Having corrected our fluxes, we compute the SFR for each spaxel utilising the Kennicutt (1998) relationship, as discussed in section 1.4.5.

Thus, the total SF traced in each galaxy by pure SF spaxels can be calculated by summing all spaxels. For our showcase galaxy, panel (j) of figure 6.3 shows the spatially resolved SFR(Hα) map.

For our "quiescent" and AGN spaxels, we combine our spatially resolved D_n4000 maps, shown in panel (g) of figure 6.3 for our example galaxy, with the relationship from Bluck et. al (2019) for D_n4000, shown in panel (i) as the solid black line, with the 1σ errors shown as the dashed black lines. The blue points show the locations of our SF spaxels on the diagram, showing agreement with the relationship. Combining these two sets of information, we recover an sSFR map of our galaxy, from which we in turn recover the SFR(D_n4000) map, shown in panel (k) of figure 6.3, by combining the sSFR map with the stellar mass map (SFR = sSFR $\times$ M_*).

Shown in panel (l) of figure 6.3 is our final SFR map is finally constructed by combining the SFR(Hα) and SFR(D_n4000) maps, replacing SFRs in our SFR(D_n4000) map with those of the SFR(Hα) for all SF spaxels.

6.3.2 The SFR-M_* diagram and ΔMS distributions

The positions of each of our galaxies on the SFR-M_* diagram is shown in the top panel of figure 6.4 as cyan points. We also show both the SDSS DR7 distribution of galaxies as the grey hexbins, coloured by density, and the MaNGA parent sample distribution, shown as contours again coloured by density. For both the SDSS DR7 and MaNGA parent sample, the masses and SFRs are taken from the MPA-JHU catalogue, wherein SFRs are calculated using either the emission lines, where applicable, or the D4000-sSFR relationship of Brinchmann et al. (2004), and stellar masses are based upon the photometry, following the theories of Kauffmann et al (2003) and Salim et al (2007). All values are calculated for the fibre observations, and aperture corrected to predict the totals for each galaxy. Whilst the majority of our galaxies do possess values within these catalogues, the determined values suffer from being based solely upon spectra taken from the centre of the galaxy, which may miss a large amount of SF from

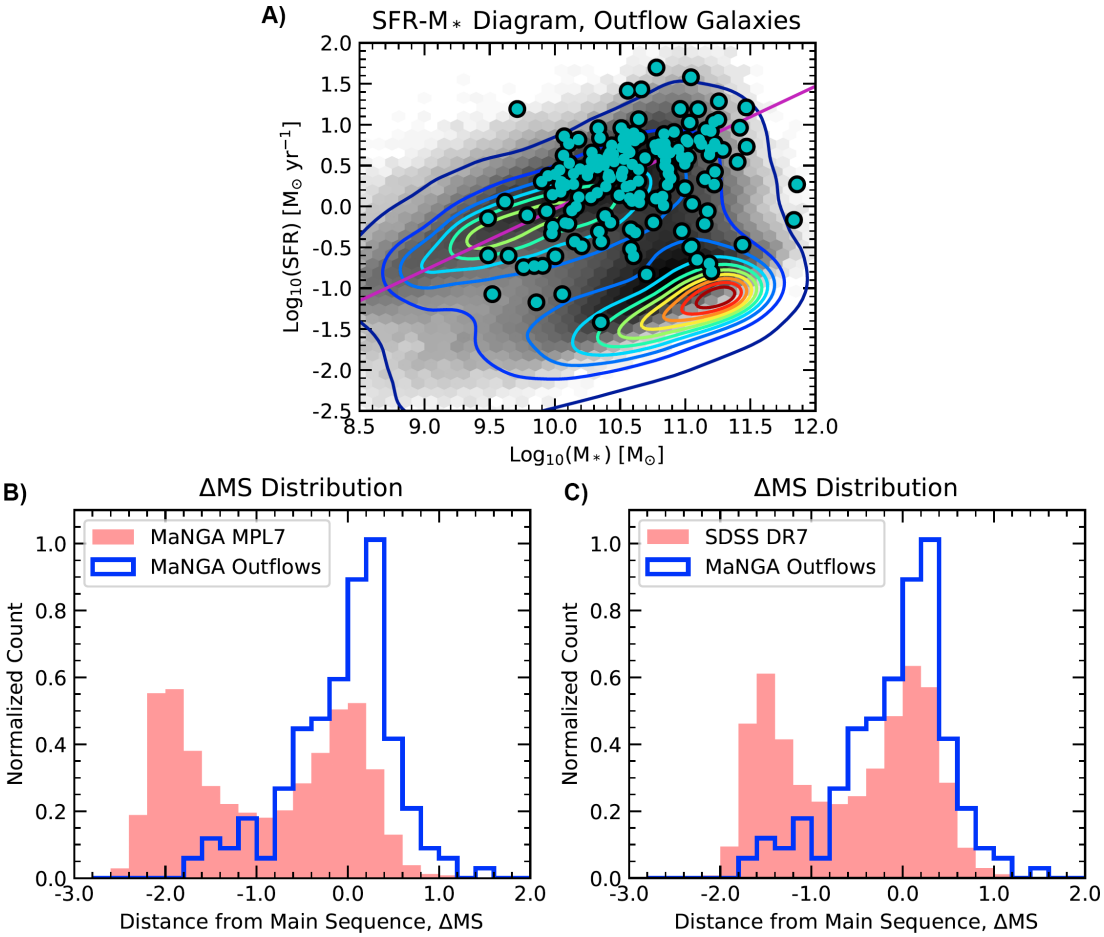


Figure 6.4: Top: SFR-M_* diagram showing the distribution of the MaNGA galaxies that host outflows. The SDSS DR7 distribution is shown in the background as grey hexbins, and the MaNGA parent sample is represented by the coloured contours. The SDSS DR7 and MaNGA parent sample values are taken from MPA-JHU catalogue, calculated from the SDSS fibre values. Bottom left: Distance from the Main Sequence (ΔMS) distributions for the MaNGA outflow sample, shown as the blue histogram, with the SDSS DR7 distribution is shown for comparison as the red histogram. Bottom right: ΔMS distributions for the MaNGA outflow sample, shown as the blue histogram, with the MaNGA parent sample distribution shown as the red histogram.

the disk, although attempts were made by Brinchmann et al. (2004) to correct for missing SF through photometric information from images. These corrections are not certain and can introduce large uncertainties, especially in the case of galaxies with low central SFRs, yet large amounts of SFR unaccounted for within the disk. These considerations lead us to take the approach discussed above for our determinations of the SFRs and M_*s for our galaxies, as opposed to simply taking the catalogue values.

We find that the majority of our galaxies hosting outflows fall on the MS, especially around

the more massive end of the diagram. Additionally, we find that a number of these galaxies fall into the starburst region, well above the MS. We also find a number of galaxies in the green valley, though not many outflows fall into the quiescent region of the diagram. The lack of low mass galaxies found in our sample may be an effect of our selection criteria, as these galaxies tend to be fainter, and thus if they hosted outflows they may be milder and fall below our detection threshold. This lack of detection alone does not serve as evidence that outflows don't exist in lower mass systems. We also find a lack of outflows in the quiescent region, though these galaxies are often early-type, hot, evolved galaxies, with large spheroidal components. Furthermore, their lack of nebular gas suitable for SF would require outflows in these galaxies to be solely AGN-driven. From our investigations into the driving sources of outflows in section 6.2.1, we can see that SF-driven outflows (as determined by the BPT diagnostics for our centred PSF selection criteria) are prevalent, accounting for between 50% and 80% of all outflows in our sample, possibly explaining the lack of outflows in the quiescent region.

We can further explore the distribution of the galaxies with respect to the MS by looking at each galaxies distance from the MS, ΔMS. This is defined purely as the distance each galaxy is from the MS on the SFR (vertical) axis. To calculate this, we need to empirically define the MS. To do this, we us the MS defined in Chang et al. (2015), wherein they studied the SFRs and stellar masses of 10^6 galaxies from a combination SDSS and WISE (Wright et al. 2010). Whilst they redefine the SFRs and stellar masses from SED fittings, they do also run their analysis on the masses derived by Brinchmann et al. (2004), which we use as our background distribution here, and define a power law that best fits the median values in mass bins of width 0.2 dex, which takes the form:

$$\log(\mathrm{SFR})\ (\mathrm{M}_\odot\ \mathrm{yr}^{-1}) = 0.75\ \log(\mathrm{M}_*/10^{10})\ (\mathrm{M}_\odot) - 0.03 \tag{6.1}$$

which can be seen in the top panel of figure 6.4, superimposed over the background distribution as the magenta line. We begin by comparing the ΔMS distribution of our outflows to that of the SDSS data, shown as the filled red histogram in panel (c) of figure 6.4. The bimodality of the SFR-M_* diagram is once again apparent from the two distinct peaks of the ΔMS distribution, one peak falling near to 0 offset (galaxies on or near to the MS) and one peak falling around an offset of -1.5, representing the quiescent galaxies. Our outflow hosts instead are predominantly on the MS. However, a tail of our outflow host galaxies fall near the peak of the quiescent galaxies from the SDSS distribution. Further, we also find a stronger representation of starburst galaxies in our sample, compared to the distribution of the SDSS data, with the peak of the outflow hosts located just above the MS.

We can also consider the distribution in comparison to the MaNGA parent sample, from which our galaxies were chosen. This distribution, shown as the red histogram in panel (b)

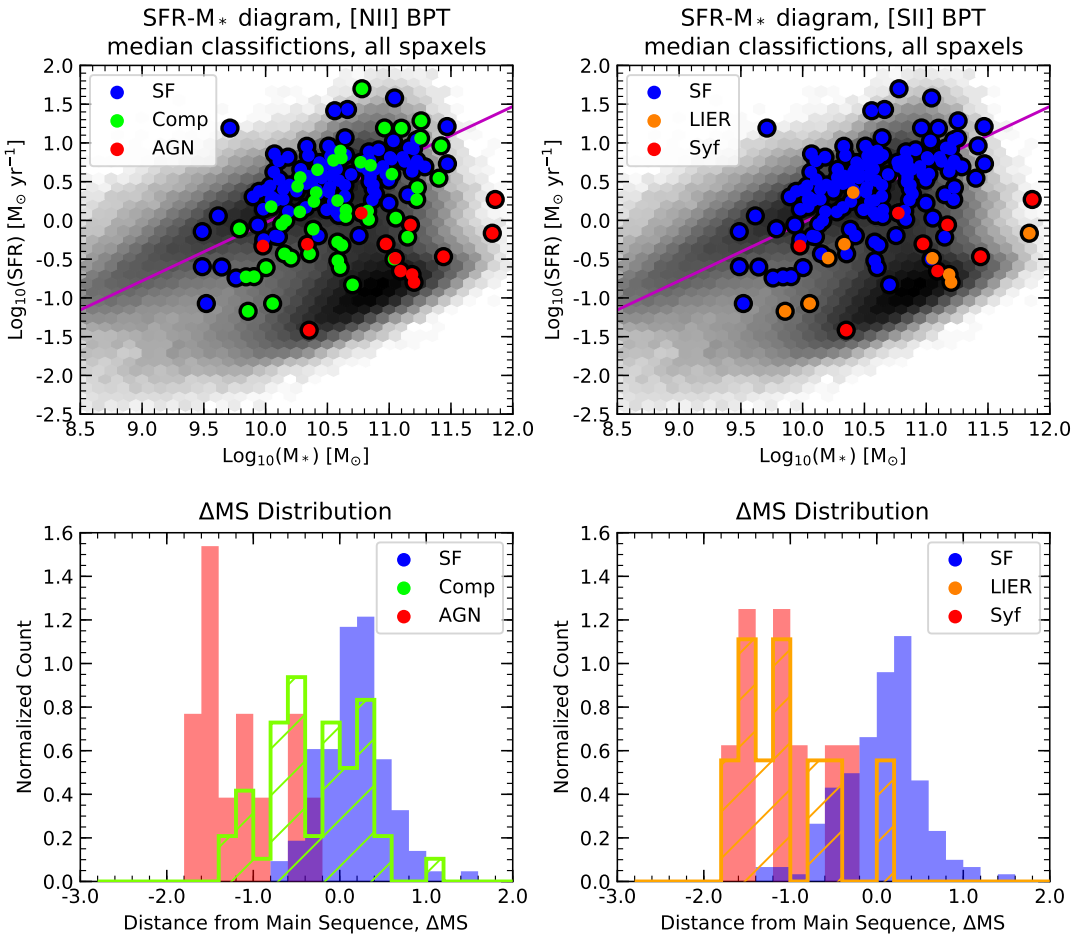


Figure 6.5: Top left and top right: SFR-M_* diagram with each galaxy point being coloured by their median [NII]- and [SII]-BPT classification for all spaxels. The black hexbins show the SDSS DR7 distribution, and the MS is represented by the magenta line. Bottom left and bottom right: The ΔMS distributions for each of the populations from the [NII] and [SII]-BPT diagrams respectively. We can see that SF outflow hosts generally fall on the MS, whilst AGN/Sy outflow hosts peak in the quiescent region.

of figure 6.4, shows a similar peak as the SDSS DR7 data near the MS, but shows a greater representation of quiescent galaxies, with the second peak falling around ΔMS=-2. Despite the MaNGA sample being drawn from the SDSS catalogue, this difference between the disturbutions can be explained by the selection requirement for MaNGA galaxies to have a flat distribution in both colour and mass, thus representing more quiescent galaxies at higher masses.

When comparing the MaNGA parent sample ΔMS distribution with that of our outflow host galaxies, the overrepresentation of MS galaxies becomes even more apparent. Once again, though it would be easy to draw the conclusion that outflows are mostly situated in galaxies on the MS, we note that our selection requirements can mean weak outflows, traced by nebular

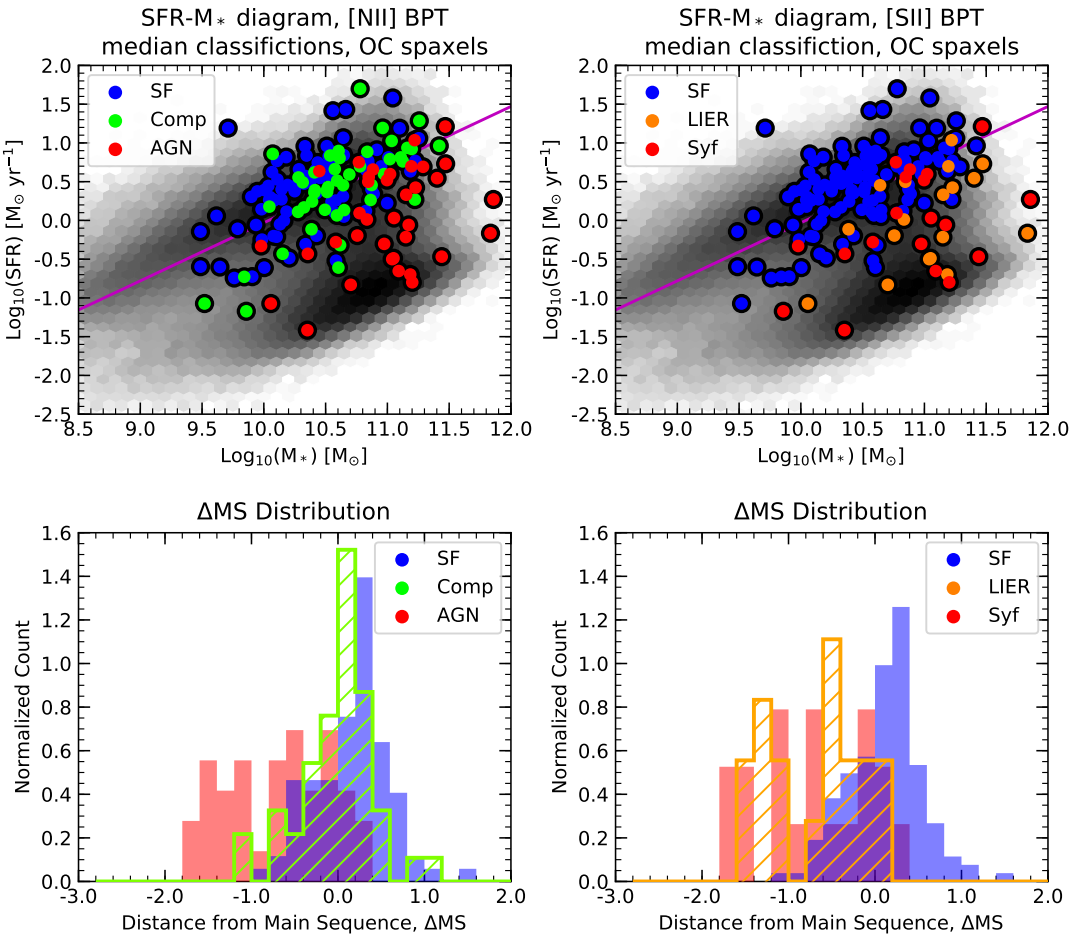


Figure 6.6: Top left and top right: SFR-M$_*$ diagram with each galaxy point being coloured by their [NII] and [SII] median BPT classification for our OC spaxels, as defined in the text and tracing the outflow driving source. The black hexbins show the SDSS DR7 distribution, and the MS is represented by the magenta line. Bottom left and bottom right: The ΔMS distributions for each of the populations from the [NII] and [SII]-BPT diagrams respectively. Here, we see that SF-driven outflows are often found in galaxies above the MS in the starburst region, suggesting the importance of starbursts for SF-driven outflows. The AGN/Sy-driven outflows are found to be more widespread, suggesting the efficiency of AGN at driving outflows in all types of galaxy.

emission lines, may be discarded. Further, as quiescent galaxies have very weak emission line strengths, outflows in these galaxies may have to be traced by other methods.

We can further consider the ΔMS distributions for our galaxies when split according to each galaxies median location on the BPT diagrams. We start with the classifications for each galaxy determined by the median location of our AS spaxels. The top left and top right panels of figure 6.5 each show the SFR-M$_*$ diagram with the points coloured by their [NII]- and [SII]-BPT classifications, respectively, whilst the bottom two panels show the ΔMS

distributions for each population. Starting with the [NII]-BPT classifications, we can see that the distributions do show a bimodality, with SF galaxies almost exclusively residing on or near the MS, Seyfert galaxies residing in the quiescent region and Composite galaxies bridging the gap in the green valley region. A similar trend can be seen for the [SII] diagram, but with SF dominated galaxies extending down into the quiescent region, suggesting that some galaxies that are in the traditional quiescent region may still be experiencing SF. This could be galaxies approaching quiescence, using the last of their gas to form stars, or galaxies in the process of being rejuvenated and not yet having moved up into the green valley. When looking at the AGN and LIER dominated galaxies, we see that most fall into the quiescent region, though their distributions do extend into the region on the diagram associated with green valley galaxies.

Next, we consider the locations instead colour coded by the driving source of the outflow, determined by them median BPT classification of the OC spaxels. Starting with the [NII]-BPT classifications, we can see that the distribution of SF-driven outflows shows a strong single peak centred around the MS. The distribution of AGN-driven outflows, however, shows a much flatter distribution with no distinct peak, still falling mainly below the MS, suggesting that the presence of AGN-driven outflows are not constrained to transitional (green valley) and quiescent galaxies, as suggested by their general galaxy classifications. Composite classifications are also again spread over both the MS and quiescent regions, though the peak here falls around the MS. This can be interpreted in a number of ways, depending on how one construes the types of ionisation responsible for a galaxy to be classified as Composite, as a Composite classification could indicate the outflow is either AGN- or SF-driven. Hence we again turn to the [SII] BPT classifications. Starting with the SF-driven outflows, we again see a strong peak around the MS, with almost all galaxies falling above the MS being classified as SF-driven, suggesting the outflow driving sources previously classified as Composite are in fact SF-driven, and that SF-driven winds are more often driven in the presence of a starburst. Likewise, considering those outflows classified as AGN-driven, we see again that, when simply considering the driving source, AGN contribution seems more widespread across the diagram, though the AGN now peak more closely to the MS. This may indicate that AGN are an important outflow driving mechanism, but their effectiveness in driving outflows is increased when they occur either on or near to the MS, likely due to presence of more ISM gas available to couple with the AGN nuclear wind and with its radiation pressure. Further, we find that a number of the galaxies around the quiescent region, with outflow driving sources classified as Composite by the [NII]-BPT, are reclassified as AGN-driven here. Finally, sources classified as LIER-driven in the [SII]-BPT diagram here are likely to be evidence of weak AGN or shock ionisation as opposed to pAGB ionisation, as we require a powerful source of energy to drive our outflows.

6.4 Gas densities in outflows and their host galaxies

6.4.1 Distributions of gas densities

The spectral coverage offered by the BOSS spectrographs ensures that, even out to the maximum redshift of galaxies in the MaNGA catalogue of z=0.15, the [SII]$\lambda\lambda$6718,6733 doublet will always be observable. As discussed in chapter 1, the electron density of the gas can be estimated from the intensity ratios of the two peaks in the doublet, as prescribed by Osterbrock & Ferland (2006). As such, we made use of the functionality of SCRAPPy to constrain the line ratios of doublets to within given limits to constrain the intensity ratio of the [SII] doublet to fall within ratios associated with densities between 40 cm^{-3} < n_e < 10,000 cm^{-3}, typical of ionised phase outflows.

We begin by investigating the typical densities found in gas in both the discs and the outflows. The top panel in figure 6.7 shows two histograms, as well as Gaussian fits to best describe their shape, of the electron densities for the disk and outflowing gas, recovered for all spaxels for which we recover both [SII] lines at a SNR>3. Interestingly, both distributions are well represented by Gaussians with widths of 0.5 dex, with the narrow (disk) component peaking around n_e = 98 cm^{-3} (log(n_e)=1.99) and the broad (outflow) component peaking approximately 0.5 dex higher in log space at n_e = 295 cm^{-3} (log(n_e)=2.47). These results are similar to others in literature that have explored the densities within outflowing gas on a spatially resolved scale (Kakkad et al. 2018, Venturi et al. 2018, Mingozzi et al. 2019), and show that the gas contained within outflows in the MaNGA survey are typically denser than the gas contained within the disk by a factor of ∼3. This is likely a consequence of the gas within the outflows being subject to higher compression due to the driving source of the outflow (Decataldo et al., 2019). However, the densities for the outflow gas are lower than other studies on outflows in the literature (Villar Martín et al. 2014, 2015, Perna et al. 2017, Förster Schreiber et al. 2019), with values typically around n_e ∼ 1000 cm^{-3}. This discrepancy was noted and discussed in Mingozzi et al. 2019, whose average densities for both the disk gas (n_e = 130 cm^{-3}) and the outflowing gas (n_e = 250 cm^{-3}) are similar to ours. They found that weighting their densities by the luminosity of each spaxel lead to density estimates similar to the literature. This suggests bias in previous studies due to consideration of only the most luminous outflow regions, characterised by high SNRs, and that high-z outflows may be more extended than the observed values suggest. An additional explanation can be prescribed to the fact that previous studies of outflows, especially high-z outflows, have targeted more extreme galaxies (such as LIRG- and ULIRG-like galaxies) which are characterised by denser outflows.

To explore effect outflows may have on the densities of gas in their host, we instead consider

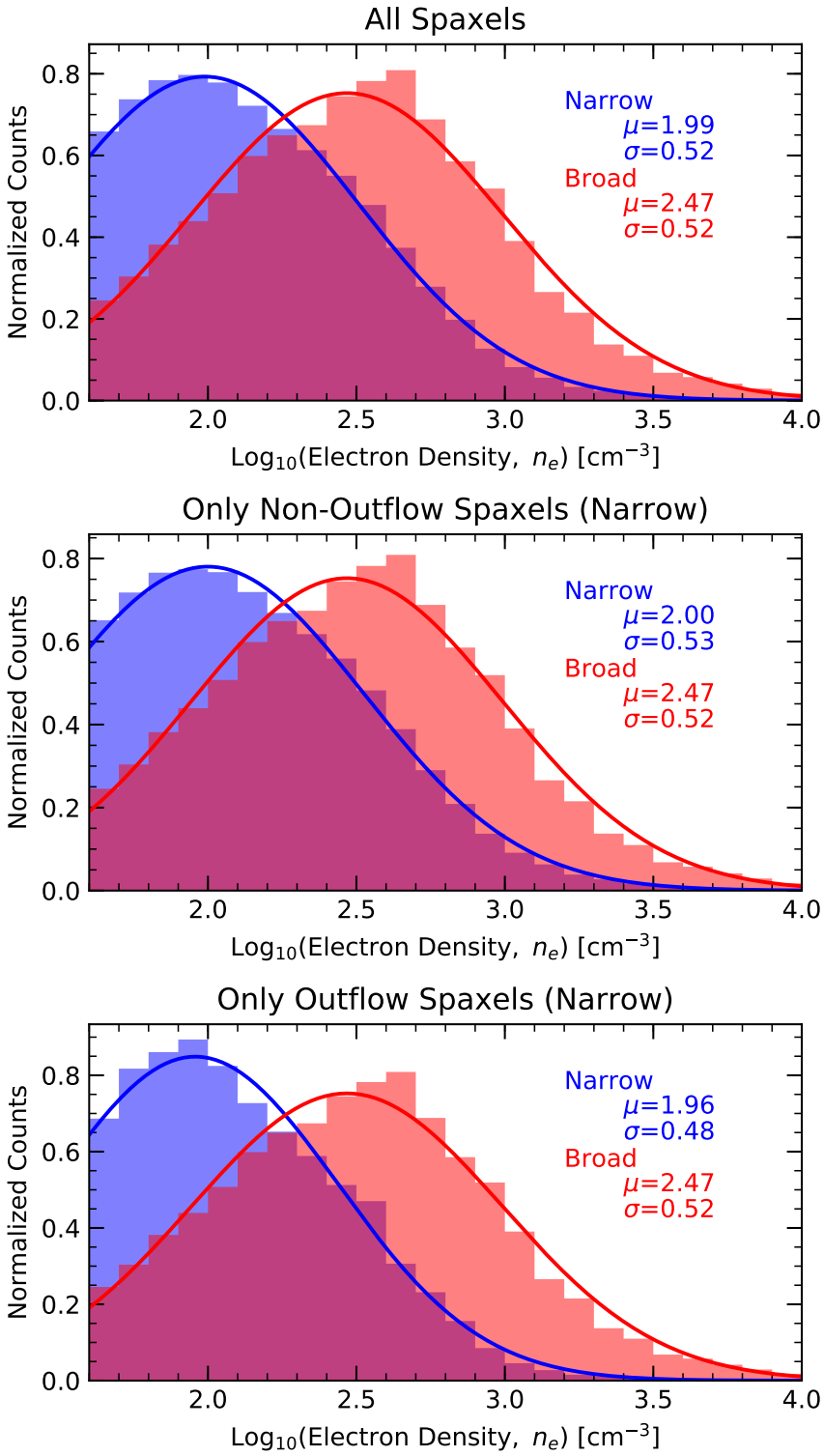


Figure 6.7: Top panel: The density distributions of the narrow (blue histogram) and broad (red histogram) component gas, as inferred from the [SII] doublet. For the narrow component, all valid spaxels are considered. Middle panel: The density distribution of the narrow component gas only in spaxels where a broad (outflow) component is also detected (blue histogram). The broad component (red histogram) remains unchanged from the top panel. Bottom panel: The density distribution of the narrow component gas only in spaxels where no broad (outflow) component is traced (blue histogram).The broad component (red histogram) remains unchanged from the top panel. From the Gaussian fits to each of the narrow component densities, we postulate that the density in the gas is unaffected by the presence of an outflow.

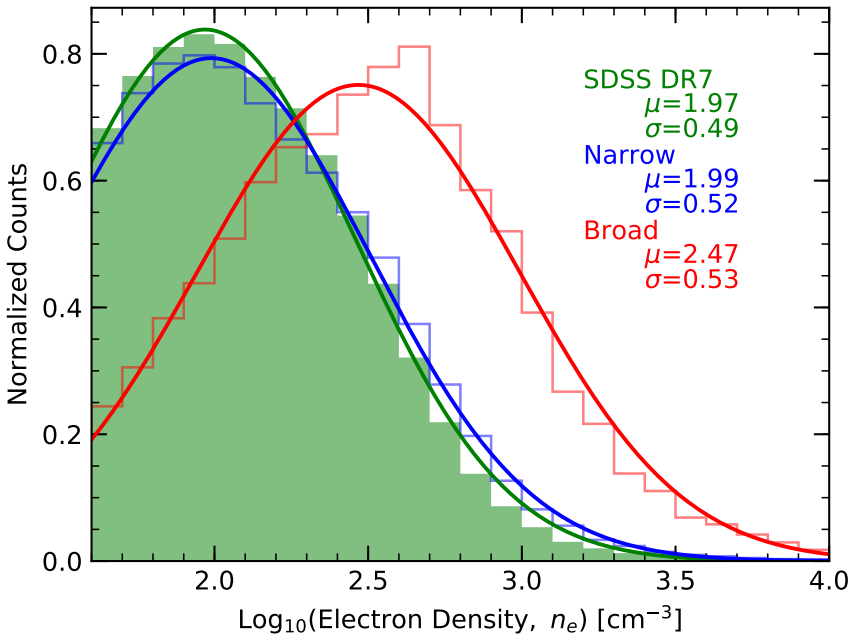


Figure 6.8: The narrow and broad component density distributions from the top panel of figure 6.7 (blue and red step histograms, respectively), compared with the density distribution of galaxies in the SDSS-DR7 dataset. This shows that the density distribution of our narrow component is similar to that of the general local galaxy population, supporting the postulate that the outflows have little to no effect on the density in the disk of their hosts.

the densities of the gas contained within two sets of spaxels: those in the disk (i.e the narrow component) that have no detection of outflows, and thus should be unaffected by the outflowing gas, and those that only show a detection of outflows. The distributions for these two sets of spaxels are shown in the middle and bottom panels of figure 6.7, respectively. Starting with the set of spaxels with no outflow detection, we can see that the electron density distribution is almost the same as the distribution for all spaxels, with only ~0.01 dex difference between the distributions means and widths. Considering instead only those spaxels with an outflow detection, we again see a similar distribution, but this time the mean and width of the distribution are both ~0.03 dex lower than those for all spaxels. The levels of differences seen are effectively insignificant, and thus we conclude that, in the MaNGA survey, outflows do not seem to have any direct effect on the gas density in the disk of their host galaxies.

One consideration here is that we are only comparing the densities in outflow host galaxies, and thus the narrow component density distribution could be affected by the outflow. To test whether our recovered disk densities are affected in any way by the presence of the outflow, we compare the total distribution for our galaxies with the distribution from the SDSS DR7 data

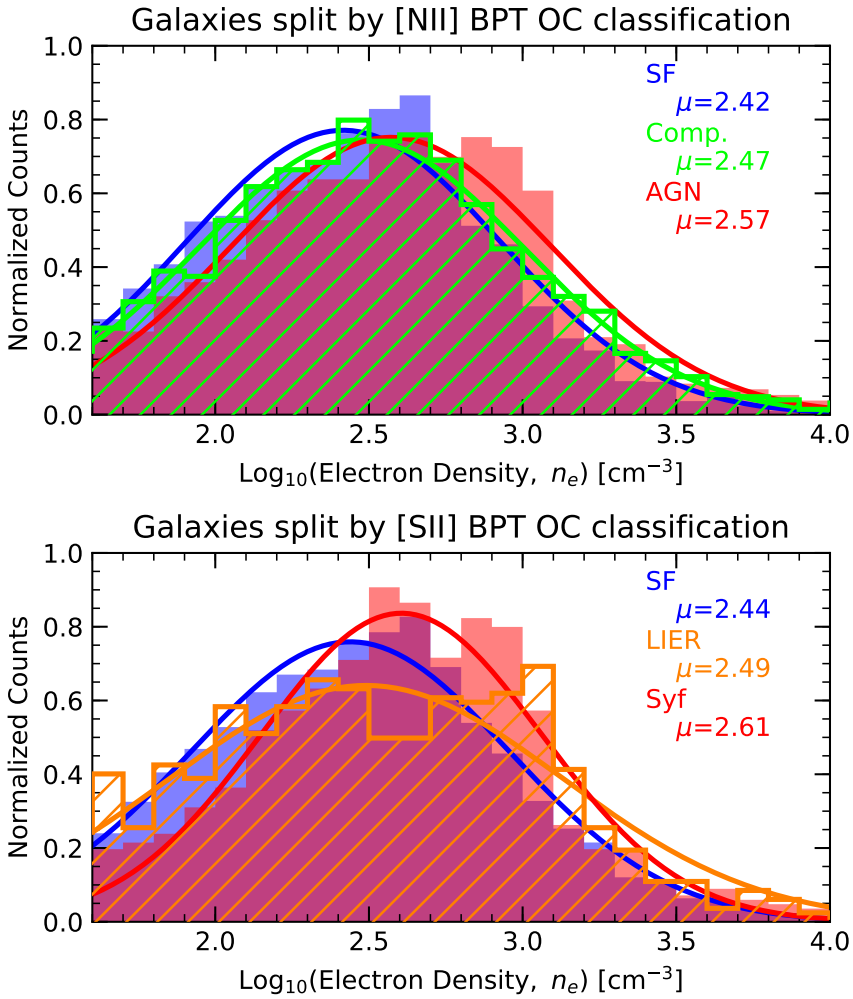


Figure 6.9: Density distributions of the broad (outflow) component gas, split by the outflow driving source in the [NII]- and [SII]-BPT diagrams (top and bottom panels, respectively). These distributions suggest that SF-driven outflows are, on average, slightly less dense than those of AGN/Sy-driven outflows.

used previously for the BPT comparisons. This comparison is shown in figure 6.8, with the DR7 distribution shown as the green filled histogram, whilst the distributions for all spaxels from our galaxies (as shown in the top panel in figure 6.7) are shown as the blue and red step histograms. From this, we can see that the median density for the SDSS DR7 falls at n_e = 93 cm^{-3} (log(n_e)=1.97), with a spread of 0.49 dex in log space, showing that the narrow component distribution of our data is consistent with the typical densities found in galaxies within the local Universe. This not only suggests that our outflow galaxy sample is not biased, but also further confirms the postulate put forth that outflows do not affect the densities in the disk in a significant manner.

Finally, we consider the differences in the density of the gas entrained within the outflows dependent on their driving source. To achieve this, we again make use of the [NII]- and [SII]-BPT median classifications of our OC spaxels, splitting our galaxies according to their classifications. The results of this division are shown in figure 6.9. Starting with the [NII]-BPT classifications, it can be seen that, to a minor extent, SF-driven outflows show lower densities than those with driving sources classified as either Composite or AGN, with AGN-driven outflows exhibiting the highest gas densities. The intermediary Composite classified driving sources show a density distribution almost identical to that of all outflow galaxies, as shown in figure 6.7. These observations are mirrored in the [SII]-BPT classification distributions, with SF-driven outflows showing the lowest average gas densities and AGN-driven outflows showing the highest average densities. The outflows with a driving source classified as LIER straddle the two other classifications, in a similar way to the Composite-driven classified outflows in the [NII]-BPT split, with a density distribution similar to the general outflow density distribution. These results suggest that, on average, AGN-driven outflows are likely to be denser than their SF-driven counterparts, in agreement with previous observations.

6.4.2 Radial gas density profiles

We next investigate how our gas densities vary as a function of radius for both components. As we wish to compare galaxies of different sizes, we normalise the radii in units of each galaxy's effective radius, r_e, for these comparisons. For the narrow component, we take the values for each galaxy from the DRP galaxy output file, which contains the parameters of an elliptical petrosian fit to photometric data, providing us with the effective radius, in arcsecs, the axis ratio, b/a, and the position angle of the major axis. For our outflow component, we take a simple radial distance from the luminosity centroid, and scale it to the galaxy's effective radius. Shown in the top two panels of figure 6.10 as black points are the individual computed densities as a function of their radius for both the disk (top panel) and outflow (middle panel) gas. Additionally, for each galaxy we calculate the average density in bins of width 0.25 r_e, plotting the average trend over the extent of the galaxy as solid blue and solid red lines for the disk and outflowing gas, respectively. Finally, for both components we show the average global trend line, calculated from the average of the individual trend lines, as a magenta line with error bars. We have restricted the analysis to be within the central 1.5 r_e as, especially for the outflow component, we have significantly less data past this radius, hence trend lines become less certain.

Both global trends show a similar decrease in density as a function of radius, with most of the decrease occurring within the central 1 r_e. The bottom panel of figure 6.10 shows both global trends together, which further exhibits the similarity in shape of the decrease in

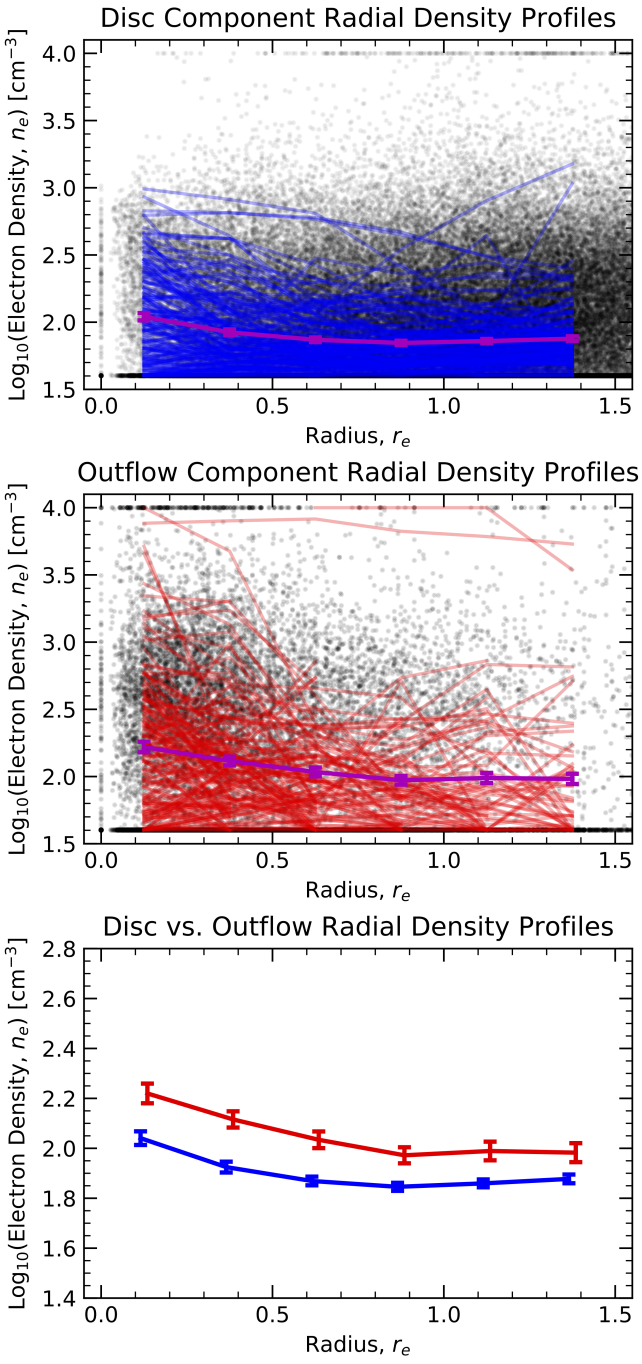


Figure 6.10: Radial density trends for the disk (narrow) component gas and broad (outflow) component gas (top and middle panels, respectively). In both panels, the individual points show the densities of all spaxels as a function of radius. The solid lines show the average trend in the density, as a function of radius, for the individual galaxies. The magenta lines show the global average trend lines, determined from the individual trend lines. The bottom panel compares the global average outflow (red) and disk (blue) trend lines from the above panels. We can see a radial decrease in density in both gas components, though the outflow component remains denser at all radii.

density. For extended [HII] regions in galactic discs, which will dominate our analysis of the densities due to the minimum SNR cut of 3 required on both [SII] lines, the trend of a radially decreasing electron density has been previously observed (Gutierrez & Beckman, 2010). The flattening of our radial trend past $\sim 1 r_e$ is likely due to the limited ability to determine electron densities via the [SII] density diagnostic below densities of 40 cm^{-3}, due to saturation of the intensity ratio. The global trend for the outflows, on the other hand, varies from what would be expected from theories in which the outflowing gas cools and fragments within the hot phase, thus forming denser clumps and molecular clouds along the direction of the outflow (Zubovas et al. 2013; Richings et al., 2018). In this scenario, we would instead expect the outflowing gas density to increase with radius. This may be a consequence of the limited spatial resolution of MaNGA combined with our large radial bins. Higher angular resolution observations are thus needed to resolve the density structure in the innermost regions of the outflows. The observed decrease of the gas densities in the outflow with radius can, however, be explained in terms of photoevaporation. Indeed, it has been noted that photoevaporation is an important effect is the regulation of the mass of molecular clumps in outflows launched by AGN (Decataldo et al. 2017, 2019). As we have shown in section 6.2.1, many of our outflows are AGN-driven, and thus may be subject to both far-UV (FUV) radiation and and extreme ultraviolet (EUV) field, leading to a fast photoevaporative flow of ionized gas.

To test this, we split our galaxies by the outflow driving source, as previously defined. The result of splitting our galaxies by outflow driving source can be seen in figures 6.11 and 6.12. The former of these figures shows the split defined by either the [NII]-BPT (left panels: SF top, Comp. middle and AGN bottom) or the [SII]-BPT (right panels: SF top, LIER middle and Sy bottom), with the average radial trends for each galaxy shown as lines, whose colour schemes remain consistent with previous figures, and the global average trend shown in each panel as a magenta line. Figure 6.12 shows each of the global average trend lines, again with a colour scheme consistent with previous definitions.

We begin by considering the AGN/Sy-driven outflows for both diagnostics. We see that in both cases, the global average trend suggests that AGN-driven outflows are the densest outflows, with a radially decreasing trend that is slightly steeper than that of the SF-driven outflows. For the [NII] diagnostic classifications, the composite global trend line does show densities similar or slightly greater than the AGN global trend line past $\sim 0.75 r_e$, though as this classification can be highly contaminated by a mixture of ongoing SF and AGN ionisation contributions, a number of the galaxies contributing to the trend line here may in fact be AGN, and as such may be classified as AGN by the [SII]-BPT diagnostic, as previously discussed in section 6.2.1. This fairly rapidly decreasing radial profile, decreasing from densities ~ 300 cm^{-3} to ~ 100 cm^{-3} within $1 r_e$, with some flattening past this radius, could be evidence for the photoevaporative flow

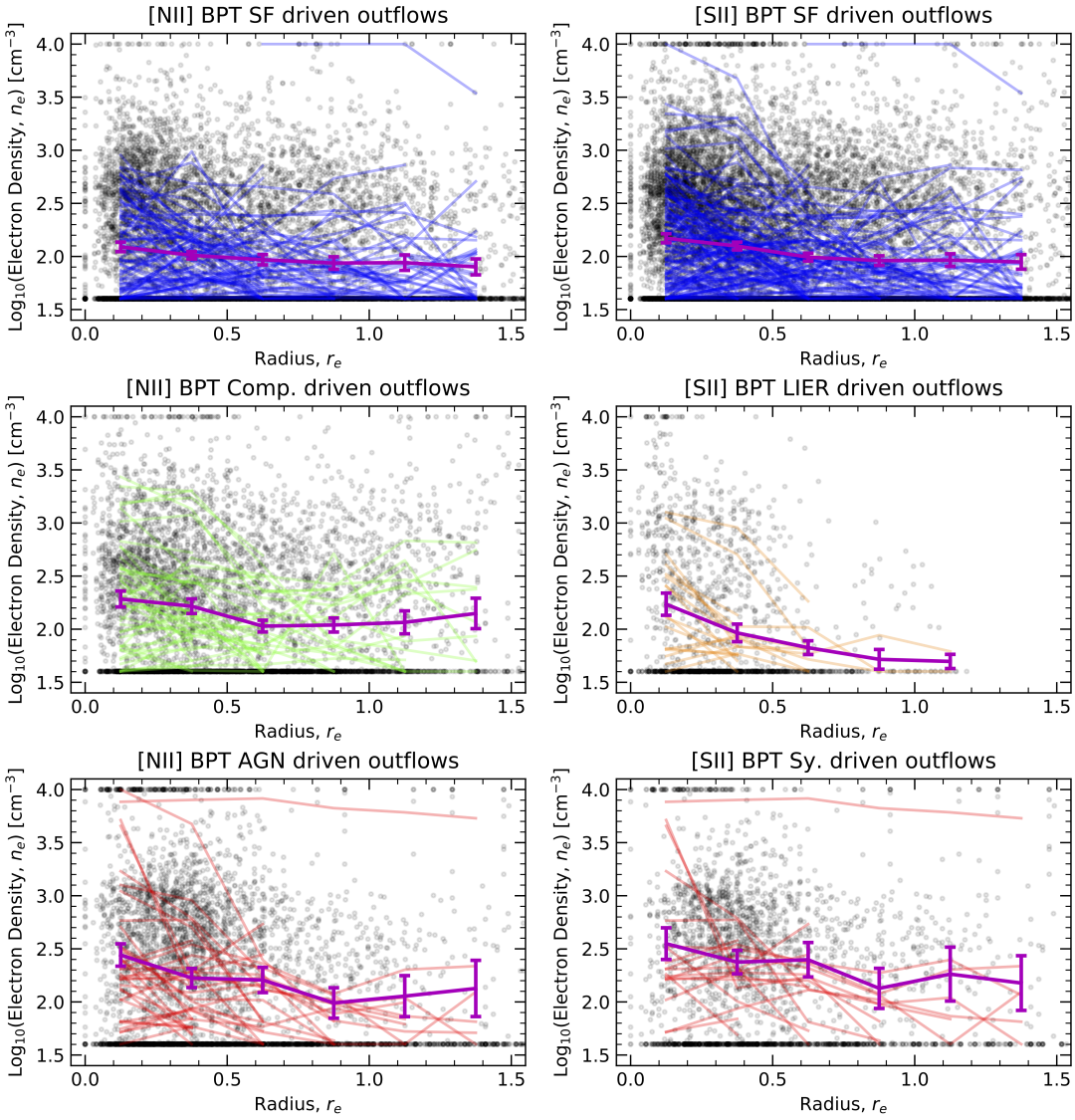


Figure 6.11: The radial density trends for the outflow component, split by their outflow driving source as classified by the [NII] and [SII]-BPT diagnostics (left and right hand panels, respectively). The colour schemes in each panel remain consistent with the colours assigned to the classifications in previous figures. The average trend lines for each classification are shown as magenta lines in their respective panels. Similar to the general outflow density distribution when split by outflow driving source (figure 6.9), we see that the AGN/Sy-driven outflows tend to be denser at then the SF-driven outflows, a trend observed at all radii. All outflows show a radial decrease in gas density as a function of radius.

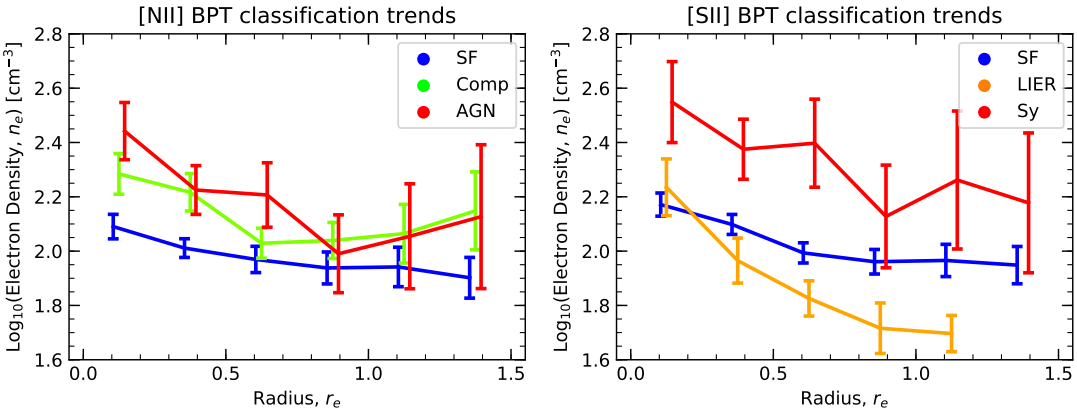


Figure 6.12: The global average trend lines of the gas density as a function of radius, split by driving source as classified by the [NII]- and [SII]-BPT diagnostics (left and right panels), from figure 6.11. The increased density at all radii of the AGN/Sy-driven outflows is better visualised here.

of ionized gas. In fact, the radiation field seen by gas within the outflow will be non-isotropic, affecting only the gas closest to the radiation source located in the nuclear region of the AGN, which could help explain the initial steep decline in the density profile, before the flattening seen at greater radii. For both classifications, the SF-driven outflow global density trends show a shallowly decreasing radial profile, at lower densities than the AGN-driven outflows at all radii, though consistently greater than the radial trend shown for the narrow disk component. Similar to the disk component, the profiles begin to flatten past $\sim 1 r_e$, likely again due to the restrictions of density calculations from the [SII] doublet.

6.5 Outflow velocities, extents and dynamical timescales

We now begin to explore the properties of our outflows, starting with the distributions of the outflow velocities, v_{outfl}, outflow radial extent, R_{outfl}, and dynamical timescales, τ_{dyn}, as functions of their positions on the BPT diagrams and the SFR-M$_*$ plane. Recovery of the first two of these properties is fairly trivial, due to our spatially resolved maps of the emission line properties of the outflowing gas. In terms of projection effects arising due to the orientation of the outflowing gas relative to the line of sight, Cicone et al. (2015) showed the average correction factor for such effects is 1, thus shouldn't statistically affect our results.

We begin by defining the outflow radial extent, R_{outfl}, as the maximum distance at which outflowing gas is still sufficiently detected relative to the outflow's driving source. For our investigations, we assume the simplest outflow model framework, wherein shells of outflowing

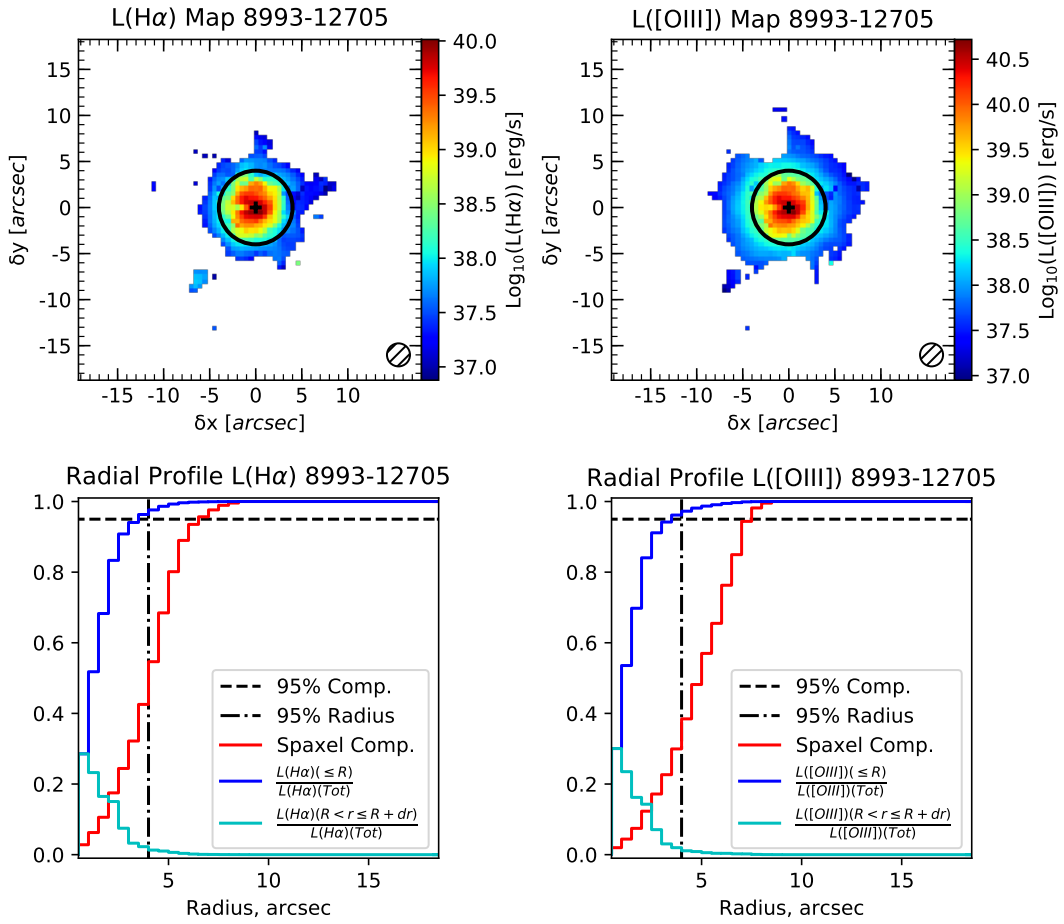


Figure 6.13: Top panels: Spatially resolved maps of the extinction corrected Hα and [OIII] luminosity of gas in the outflow. The black circles represent the outflow extent, as defined in the text, for each emission line. Bottom panels: a graphical representation of the determination of the outflow extent. The dark blue line shows the total luminosity internal to radius R, whilst the light blue line shows the radius contained within an annulus, of width dR=0.5 arcsec, between radii R and $R + dR$. The red line indicates the fraction amount of spaxels contained at each radius. The outflow extent is defined as the radius at which a minimum of 95% of the outflow luminosity, as indicated by the horizontal dashed black line, is contained. The outflow extent is shown as the vertical dot-dashed black line.

material expand spherically from a central driving source. Whilst this model is an incredibly simplistic view of the nature of outflows, it provides us a general framework from which we can compare all of our outflows without the need to fully model each outflow. As such, within this framework, we define out outflow extent in a similar fashion to how the effective radius is typically defined. Namely, we place a circular aperture over the outflow's central luminosity spaxel, expanding its radius until a sufficient amount of the outflow's total luminosity

is contained within. In choosing the total luminosity we wish to contain, we much consider a number of limitations due to the data. Specifically, we must take into account nominal outflow detections can be made in the extremities of the galaxies due to artefacts or noise spikes, and beam smearing effects of our observations due to the MaNGA spatial PSF. As such, we elect to define the extent to be the radius within which a minimum of 95% of the total outflow luminosity is contained. Figure 6.13 shows the determination of R_{outfl} for the same sample galaxy as shown in figures 6.1 and 6.3. For each galaxy, we determine the radial extent for both Hα and [OIII] separately, with the aperture centred on the luminosity centroid spaxel of each species. As the MaNGA data has a spatial resolution of 0.5 arcsec per spaxel, we increase the radius in steps of 0.5 arcsec, thus recovering R_{outfl} in units of arcsec. Using the galaxy's galactocentric redshift, we convert this distance to kpc. Whilst both values are later used in outflow rate calculations (discussed in section 6.6.2), when discussing the outflow extent we refer to the maximum of these two quantities.

Next, we determine our outflow velocities, v_{outfl}, defined as $|v|$ + FWHM/2 for the broad component, where $|v|$ is the absolute velocity of the outflow compared to the galaxy's rest frame velocity, defined by the stellar component velocity, and the FWHM is determined from the broad component velocity dispersion, σ_{broad}. To account for the projection of the outflow not necessarily being along our line of sight at the outflow's central spaxel, we consider all spaxels that fall within R_{outfl}, taking our final v_{outfl} as the 98^{th} percentile velocity in these spaxels. In this way, regardless of the outflow projection, we should recover a velocity close to the true maximum outflow velocity.

Finally, we determine the dynamical timescales of our outflows, τ_{dyn}. Having v_{outfl} and R_{outfl}, we define the dynamical timescale as the time taken for the outflowing gas to travel from the outflow centre to R_{outfl} at velocity v_{outfl}, namely:

$$\tau_{dyn} = \frac{R_{outfl}}{v_{outfl}} \tag{6.2}$$

With these three quantities calculated, we begin by considering the distribution of outflow velocities, shown in figure 6.14. The upper panels of the diagram show a large scatter, however we can tentatively see in both of the BPT diagrams that the AGN/Seyfert loci on both diagrams host outflows with greater velocities, on average, than those found in the SF loci, suggesting that AGN-driven outflows are likely faster than SF-driven outflows. This result is in agreement with expectations of models and hydrodynamical simulations, according to which SF-driven outflows can only reach velocities of the order of a few hundred km s^{-1}, whilst AGN-driven outflows can exceed 1000 km s^{-1} (e.g. Costa et al. 2015). Furthermore, in the [SII]-BPT classifications, for LIER driving sources we observe outflows with high velocities, comparable to those of AGN-driven outflows, which is probably indicative that the LIER-like emission

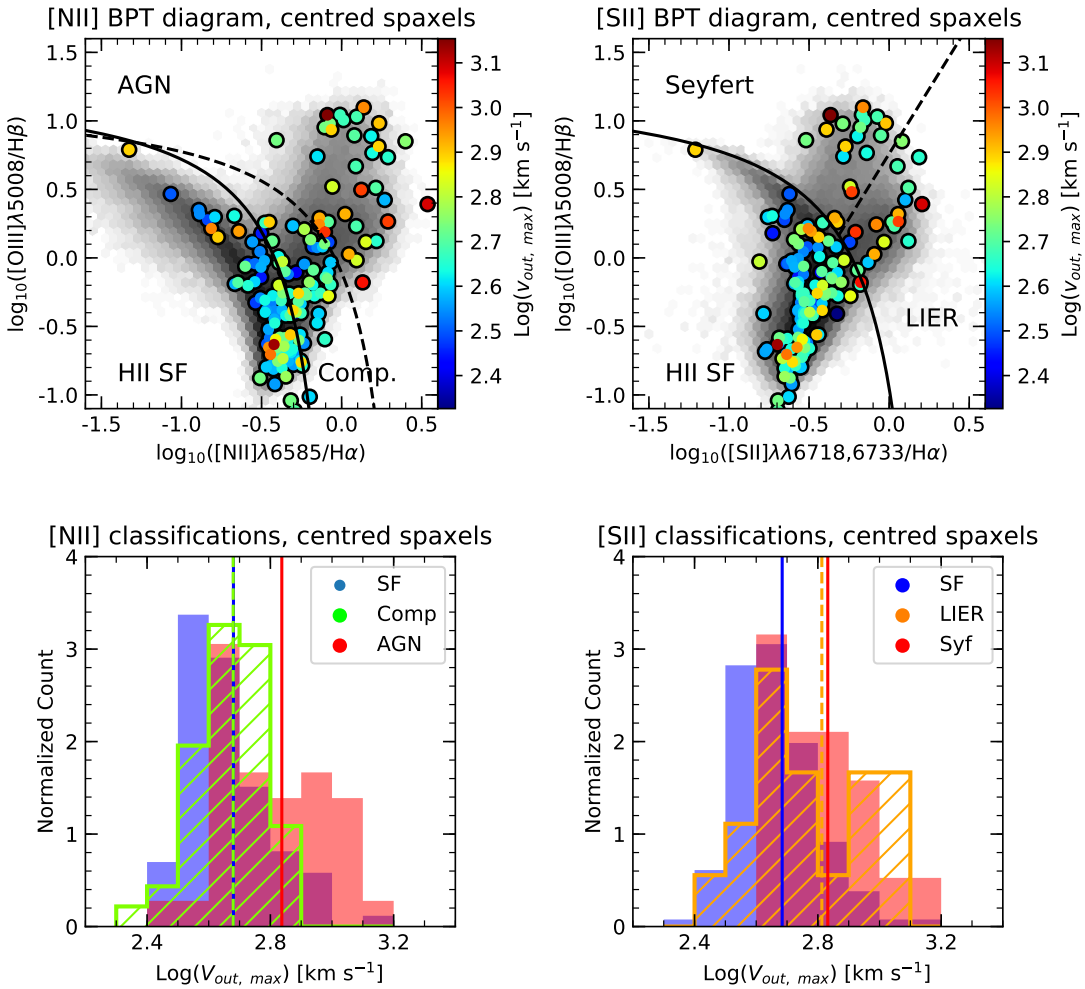
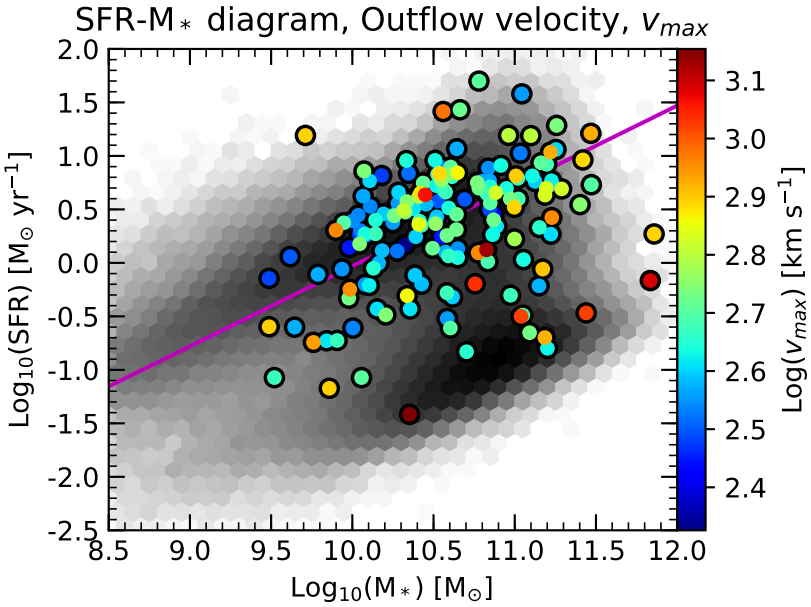
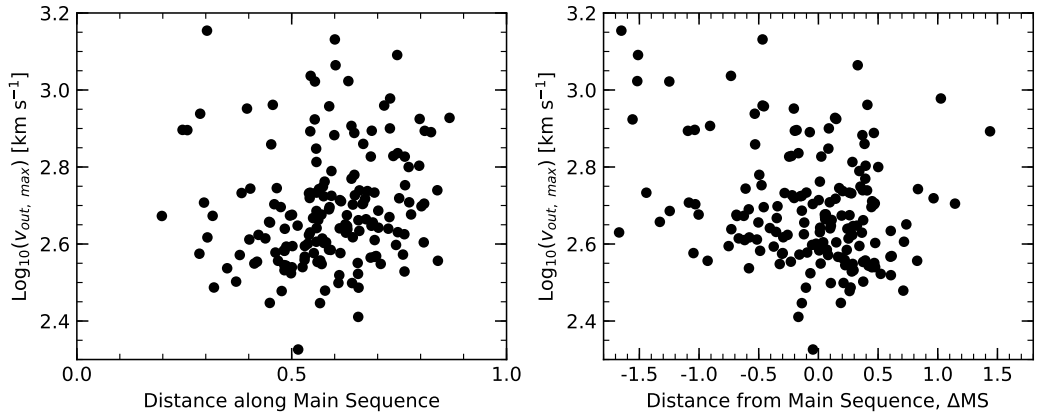


Figure 6.14: Top panels: Distribution of outflow driving sources on the [NII]- and [SII]-BPT diagrams (left and right panels, respectively), coloured by their outflow velocity, v_{outfl}. Bottom panels: Distributions of v_{outfl}, as split by their [NII]- and [SII]-BPT classified driving sources (left and right panels, respectively). We can see here that the highest velocity outflows are typically in AGN-driven outflows, reaching over 1000 km s^{-1}, whilst SF-driven outflows show the lowest velocities, rarely exceeding a few hundred km s^{-1}, in agreement with numerical models.

is due to strong shocks. However, when we consider instead the SFR-M$_*$ diagram in figure 6.15(a), we see no clear correlation between outflow velocity and location. Whilst no significant trends are seen, the lowest velocity outflows are found almost exclusively along the MS, whilst the largest velocities tend to be in galaxies located away from the MS, either in galaxies along the red sequence below the MS, or in the starburst region above it. This is likely because along the MS is where most mild SF activity (and hence a modest SNe rate) is present, weakly



(a)



(b)

Figure 6.15: Upper panel: The distribution of the outflow host galaxies on the SFR-M$_*$ plane, colour-coded by their outflow velocity, v_{outfl}. We see here than most of the low v_{outfl} galaxies are on or slightly above the MS, whilst the highest v_{outfl} outflows are found away from the MS in either the starburst region or the quiescent region (in outflows previously determined to be AGN-driven). Bottom panels: The distributions of v_{outfl} as a function of distance along the MS, as defined in the text, and ΔMS. No significant trends arise in these plots, though the higher velocity outflows are again seen to be further away from the MS in the ΔMS plot.

driving outflows, whereas the strongest starbursts (and hence the highest SNe rates) are located

above the MS and the most powerful AGN are below the MS, both more powerfully driving outflows. We explore these trends further by considering the outflow velocity as a function of both distance from the MS, ΔMS, and our newly defined distance along the MS. This quantity is measured in the dimension of the MS, with 0 and 1 being defined as the locations on the MS at stellar masses of $10^{8.5}$ and 10^{12}. The distance itself is calculated as the location of the perpendicular intersection of each galaxy and the MS. Figure 6.15(b) shows plots of v_{outfl} against these two distance measures. Whilst the plot of v_{outfl} against distance along the MS again shows no strong trends, in the v_{outfl} against ΔMS we again tentatively see that outflow velocities seem to increase with distance (both positive and negative) from the MS. Of course, as mentioned we only observe the velocities along our line of sight, so projection effects that are difficult to assess with the limited spatial resolution offered by MaNGA will also play a role.

We next consider the extent of the outflow as a function of the galaxy locations on the BPT and SFR-M_* diagrams, as shown in figures 6.16 and 6.17 respectively. Whlist it can be seen from the histograms that, on average, AGN-driven outflows are *very* slightly more extended (around 0.5 kpc in both the [NII]- and [SII]-BPT classifications) and our most extensive outflow is AGN-driven, a large fraction of the most extended outflows are observed amongst SF-driven outflows. This is most apparent in the [SII]-BPT diagram and associated histogram. In fact, many of the extended AGN-driven outflows are also located very close to the SF-AGN demarcation line, suggesting SF may indeed still play a role in driving the outflow. These extended SF-driven outflows are likely due to the fact that in SF galaxies the outflow driving sources (SNe and stellar winds) are distributed on larger scales (throughout the disk). The few AGN-driven outflows that are observed to be very extensive are likely due to powerful AGN driving these outflows to the large distances observed. When we instead consider the SFR-M_* diagram, there is a clear and distinct trend with outflow extent increasing along the MS. As we move to more massive galaxies, in turn hosting larger amounts of ongoing SF, the extent to which the outflows are driven is larger, suggestive of powerful driving sources. The most extensive of our outflows are found in the starburst region above the MS and at the massive end of the main sequence. This trend is most obvious in the bottom left panel of figure 6.17, showing the outflow extent as a function of distance along the MS. Despite the moderate scatter, there is a clear, increasing trend of outflow extent with distance along the MS. Considering instead how the extent varies as a function of ΔMS, no clear trend arises, suggesting that the increase in extent with distance along the MS is a function of both the stellar mass *and* SFR of a galaxy, not just SFR alone. As we have seen, the maximal velocity of these outflows may not be large, suggesting that although these outflows are extensive, the ionised gas may not escape from the gravitational potential of these large galaxies. Furthermore, the suggestion of large, extended outflows in galaxies in this region could go some way to supporting the suggestion that stars that may form in these

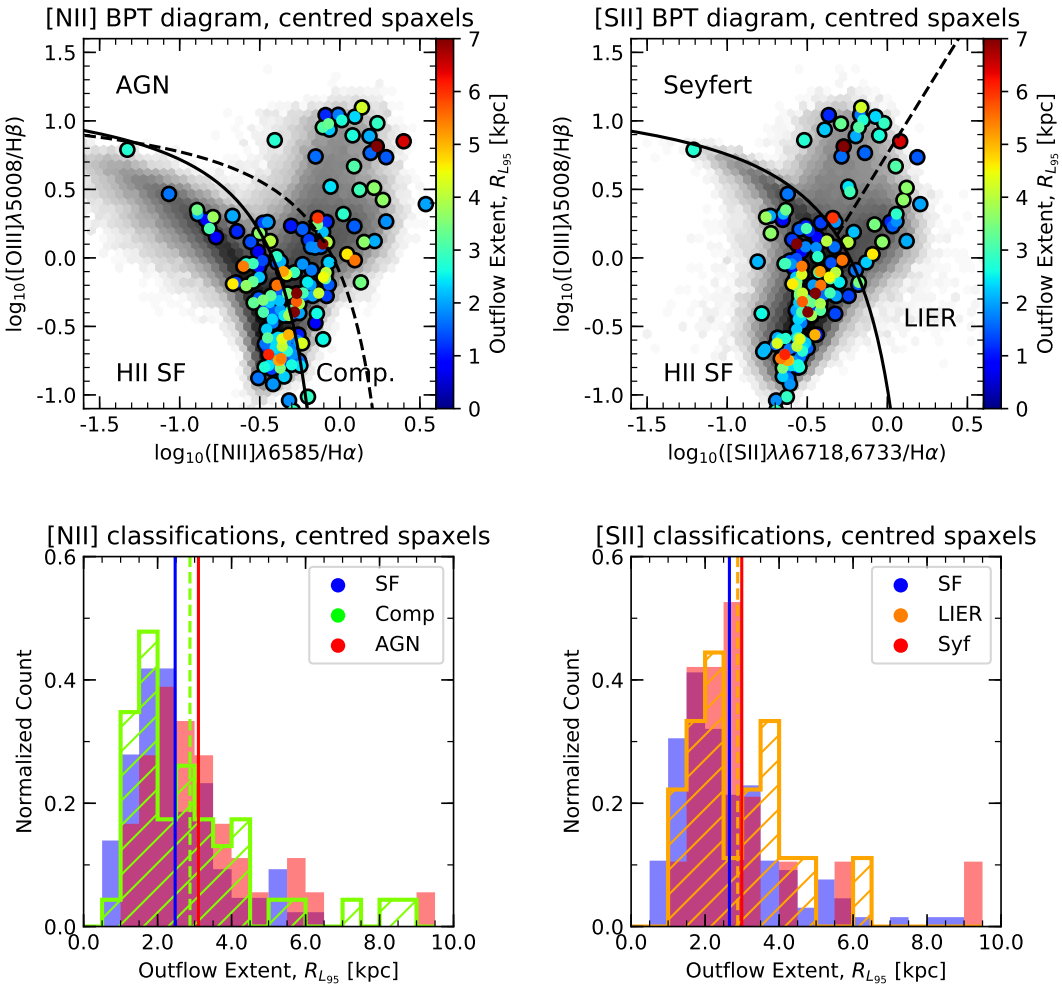
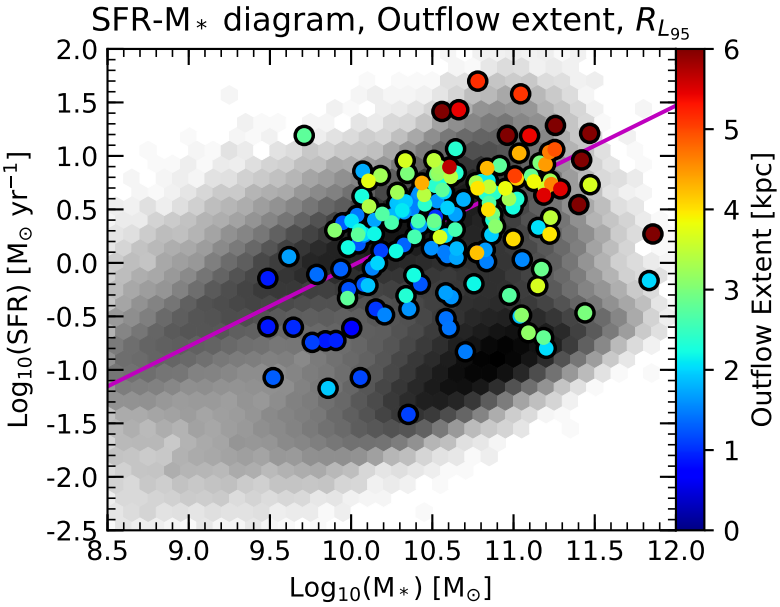


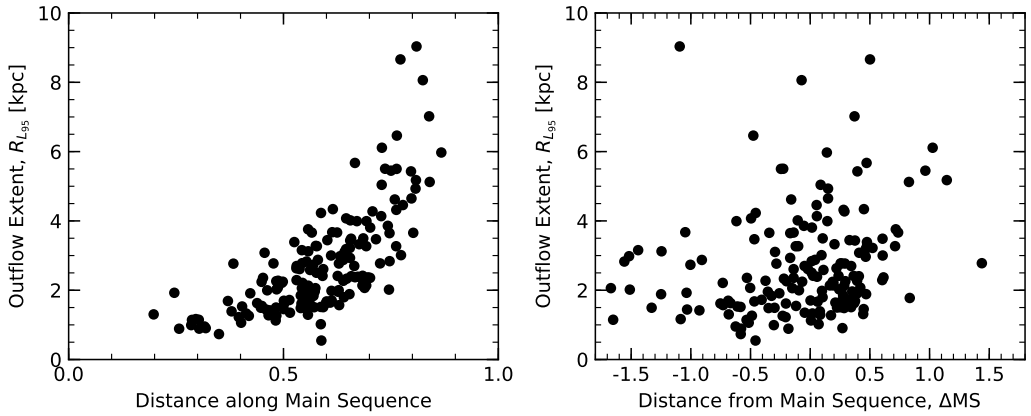
Figure 6.16: Top panels: Distribution of outflow driving sources on the [NII]- and [SII]-BPT diagrams (left and right panels, respectively), coloured by their outflow extent, R_{outfl}. Bottom panels: Distributions of R_{outfl}, as split by their [NII]- and [SII]-BPT classified driving sources (left and right panels, respectively). We can see here than, whilst AGN-driven outflows seem to be slightly more extended on average, most of the extended outflows as SF-driven, likely due to their driving sources (SNe and stellar winds) being spread over larger scales.

outflows (as discussed in chapters 4 and 5) could contribute to the bulge stellar population, or that the re-accretion of this gas could cause starburst periods, a suggestion in line with recent observations of galaxies above the MS with bluer bulges, suggesting ongoing SF (Morselli et al. 2017).

Finally, we consider the distribution of the dynamical timescale, τ_{dyn}, of our outflows as a function of the galaxy locations on the BPT and SFR-M$_*$ diagrams, shown in figures 6.18 and 6.19 respectively. As this quantity is defined by the ratio of the previous two explored



(a)



(b)

Figure 6.17: Upper panel: The distribution of the outflow host galaxies on the SFR-M$_*$ plane, colour-coded by their outflow extent, R_{outfl}. We can see here that the most extended outflows tend to be in galaxies at the upper-mass end of the diagram, and preferentially in starburst galaxies. Bottom panels: The distributions of R_{outfl} as a function of distance along the MS, as defined in the text, and ΔMS. Whilst we see no significant trend in the ΔMS plot, there is a clear correlation between increasing distance along the MS and increasing R_{outfl}. This suggests that R_{outfl} may be influenced by an increase in *both* SFR and stellar mass of the host. The propensity of these extended outflows to be in the starburst region is again likely due to the outflow driving sources being distributed on larger scales than for AGN-driven outflows.

quantites ($\tau_{dyn} = R_{outfl}/v_{outfl}$) we expect to see trends linked to those of R_{outfl} and v_{outfl}. In fact, the dynamical timescale seems to most closely follow the trends of the radial extent of the outflow, showing some correlation with position in the BPT diagrams, with typically slightly longer timescales for SF-driven outflows. This is again most obvious in the [SII]-BPT diagram and histogram, though the average timescale is (slightly) greater for SF-driven outflows in both histograms. We also see and increase in timescale with increasing distance along the MS. This trend is most readily obvious in the left panel figure 6.19(b) where, despite the moderate scatter, a linear increase in τ_{dyn} as a function of distance along the MS is obvious. These results suggest, not unexpectedly, that SF activity can drive outflows more continuously for longer timescales, especially in massive SF and starburst galaxies. On the contrary, due to their flickering nature AGN drive outflows in a more bursty mode and for shorter periods of time.

6.6 Outflow masses and outflow mass rates

6.6.1 Ionised gas outflow masses

6.6.1.1 Ionised gas outflow mass calculations

For the determination of our outflow masses, we use the same methodology as in chapter 5. However, as discussed in section 6.4, we have recovered gas densities for our outflows from the [SII] doublet. Whilst the determination of outflow masses using the spatially resolved density maps would be a novel approach to the calculations, this approach would require us to have both peaks in the [SII] doublet detected at an SNR≥3 in all spaxels for which we have either Hα or [OIII] traced at SNR≥3. Unfortunately, in our data, we only trace both [SII] peaks to this level in 39% of all good outflow spaxels, with a typical completeness of only 37%. Furthermore, some galaxies exhibit no spaxels with both peaks at SNR≥3, meaning we cannot simply use an estimate of the electron density from the highest [SII] SNR spaxel in the galaxy. As such, we have elected to use the median value of our distribution from section 6.4.1, $n_e = 295$ cm^{-3}, for the electron density in equations 1.6 and 1.7 to determine the outflowing gas mass in each spaxel from our (extinction corrected) L([OIII]) and L(Hα).

The top panels of figure 6.20 show the spatially resolved maps of the computed Hα and [OIII] outflow gas masses (M_{outfl}(Hα) and M_{outfl}([OIII]), respectively) for the same example galaxy used throughout this chapter, with the solid black circle indicating the previously defined Hα and [OIII] outflow extent radii. Also shown in the figure are the determinations of the Hα and [OIII] mass outflow rates, which are discussed in section 6.6.2.

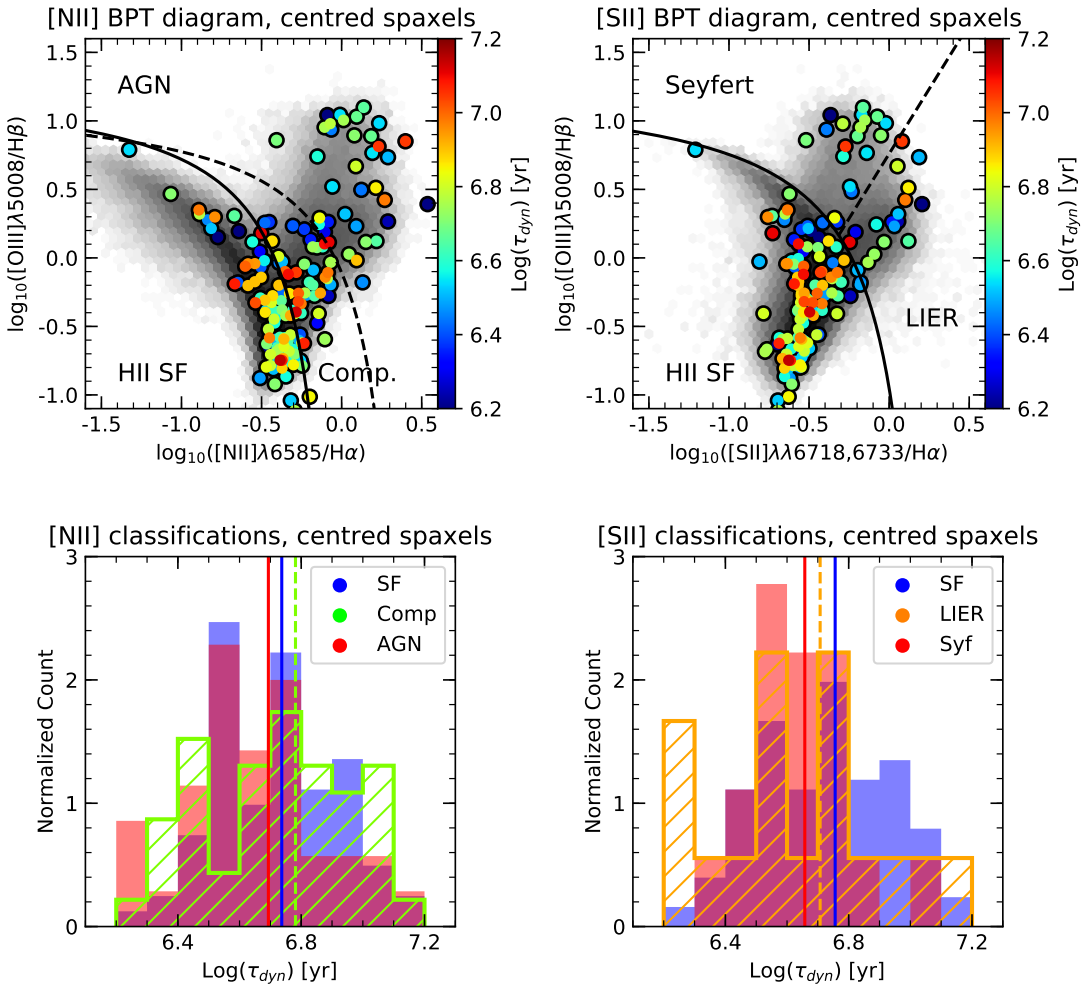
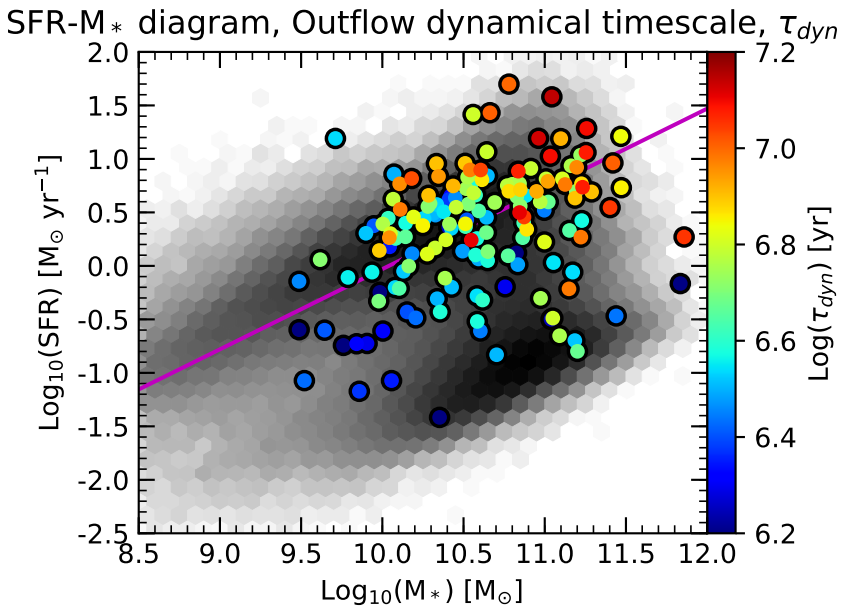


Figure 6.18: Top panels: Distribution of outflow driving sources on the [NII]- and [SII]-BPT diagrams (left and right panels, respectively), coloured by their dynamical timescale, τ_{dyn}. Bottom panels: Distributions of τ_{dyn}, as split by their [NII]- and [SII]-BPT classified driving sources (left and right panels, respectively). We can see here that SF-driven outflows seem to operate on longer timescales than AGN-driven outflows, consistent with a flickering nature of AGN.

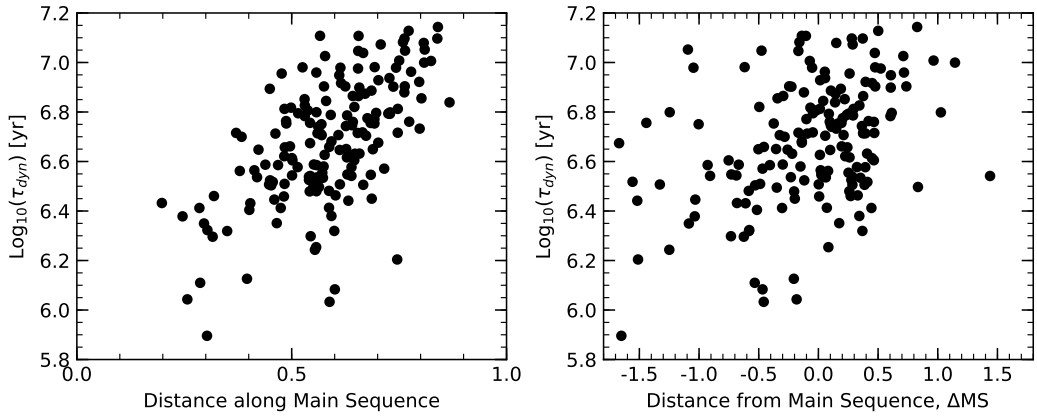
6.6.1.2 Distribution of outflow masses on the BPT-diagrams and SFR-M$_*$ plane

We now explore the distribution of the Hα and [OIII] inferred outflow masses as a function of position on the BPT and SFR-M$_*$ diagrams.

We begin by considering the distribution with respect to the outflow mass inferred from Hα, M_{outfl}(Hα). The upper panels of figure 6.21 shows the locations of our outflows on the [NII]- and [SII]-BPT diagrams, colour-coded by their M_{outfl}(Hα), with the histograms of their distributions, split by locus, shown below. As can be seen, M_{outfl}(Hα) seems to show very



(a)



(b)

Figure 6.19: Upper panel: The distribution of the outflow host galaxies on the SFR-M$_*$ plane, colour-coded by their dynamical timescale, τ_{dyn}. We see here a similar trend to R_{outfl}, with the longest operating outflows in galaxies at the upper-mass end of the diagram, and preferentially in starburst galaxies, though as τ_{dyn} is calculated from R_{outfl} (and v_{outfl} this similar trend is unsurprising. Bottom panels: The distributions of τ_{dyn} as a function of distance along the MS, as defined in the text, and ΔMS. Whilst we see no significant trend in the ΔMS plot, there is a potential correlation between increasing distance along the MS and increasing τ_{dyn}. This suggests that τ_{dyn} may be influenced by an increase in *both* SFR and stellar mass of the host, similarly to R_{outfl}. The propensity of SF-driven outflows in the starburst region to operate on longer timescales than AGN-driven outflows is again consistent with aflickering nature of AGN.

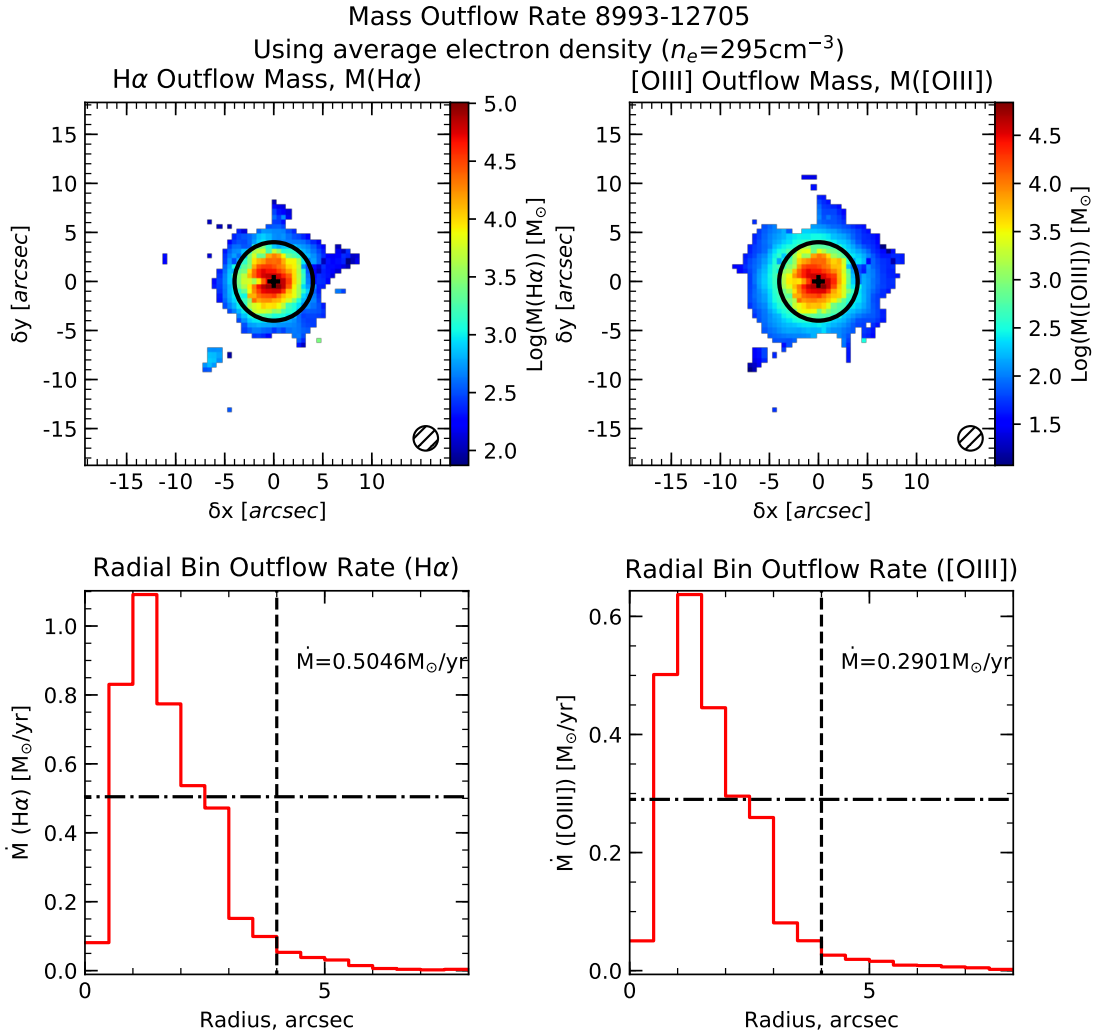


Figure 6.20: Top panels: The spatially resolved maps of the outflow mass as inferred from Hα and [OIII] (left and right panels, respectively). Also shown as the black circle is the outflow extent radius, as defined in the text. Bottom panels: a graphical representation of the outflow mass rate calculation using the radial bin methodology, as discussed in section 6.6.2.1.1. The outflow mass rate in each 0.5 arcsec width radial bin is shown by the red histogram. The outflow rate is then defined as the average of all bins internal to the outflow extent, indicated by the vertical dashed black line. The horizontal dot-dashed black line shows the average outflow mass rates inferred from Hα and [OIII] (left and right panels, respectively).

similar distributions for the SF and Composite classifications in the [NII]-BPT diagram (solid blue and hashed green histograms in the lower panels of figure 6.21), with the average outflow mass of each classification of outflow driving source begin incredibly similar (the blue solid and green dashed lines). This suggests that many of the outflows in the Composite regions

are likely SF-driven, though a number of the outflows do exhibit outflow masses more in line with those of the AGN classification distribution (red histogram). The AGN classified outflow driving sources themselves exhibit an average $M_{outfl}(\mathrm{H}\alpha)$ greater than that of the other two classifications, but only by a factor of around 2 (0.3 dex logarithmically).

Similarly, for the [SII]-BPT the SF and LIER classified outflows driving sources show distributions with similar average outflow masses (distributions shown as the blue filled and orange hashed histograms, with average values as the solid blue and dashed orange lines, for the SF and LIER sources, respectively). The Sy classified driving sources again exhibit $M_{outfl}(\mathrm{H}\alpha)$ around a factor of 2 (0.3 dex logarithmically) greater than the other classifications, with a majority of the more massive outflows being classified as Sy-driven (red histogram and red dashed line in the lower left panel of figure 6.21 showing the distribution and average outflow mass, respectively).

Moving to figure 6.23, which shows the distribution of the outflow galaxies on the SFR-M_* diagram coloured by their $M_{outfl}(\mathrm{H}\alpha)$ in the right hand panel, we see that the more massive outflows, as inferred by $M_{outfl}(\mathrm{H}\alpha)$, are found in galaxies residing at the top of the main sequence and in the starburst region just above it. We do also find some outflows with intermediate outflow masses in the red sequence, suggesting that these galaxies may be experiencing the last stages of outflows removing their gas reservoirs, thus quenching them and transitioning them to the red sequence. An alternative interpretation is that the galaxies may have recently accreted some gas, 'rejuvenating' the galaxy and beginning to move it towards the green valley, though now experiencing BH accretion and thus an AGN-driven outflow removing and heating this gas (all of these outflows are classified as AGN-driven, or exhibiting LIER-like ratios, in both the [NII]- and [SII]-BPT diagrams, shown in figure 6.5). In fact, a similar phenomenon has been suggested for the so-called 'red-geysers' (quiescent galaxies with winds driven by low-luminosity AGN heating and removing cold accreting gas) also discovered by MaNGA (Cheung et al. 2016).

Figure 6.22 shows the same as figure 6.21, but instead shows the galaxies colour coded by their outflow masses inferred from [OIII], $M_{outfl}([\mathrm{OIII}])$. In this figure, we can see that the $M_{outfl}([\mathrm{OIII}])$ of those galaxies defined to be AGN/Sy-driven in the BPT diagrams are much more massive than those in either the SF or Composite loci in the [NII]-BPT diagram and the SF or LIER loci in the [SII]-BPT diagram. When looking at the distributions for the [NII]-BPT diagram, we can see that the SF and Composite classified driving sources (blue filled and green hashed histograms, respectively) again show similar distributions of $M_{outfl}([\mathrm{OIII}])$, though a number of the Composite classified sources do possess masses more in line with the AGN distribution (red filled histogram). In agreement with the observations of $M_{outfl}(\mathrm{H}\alpha)$ this suggests most of the outflow driving sources classified as Composite are likely to be due more

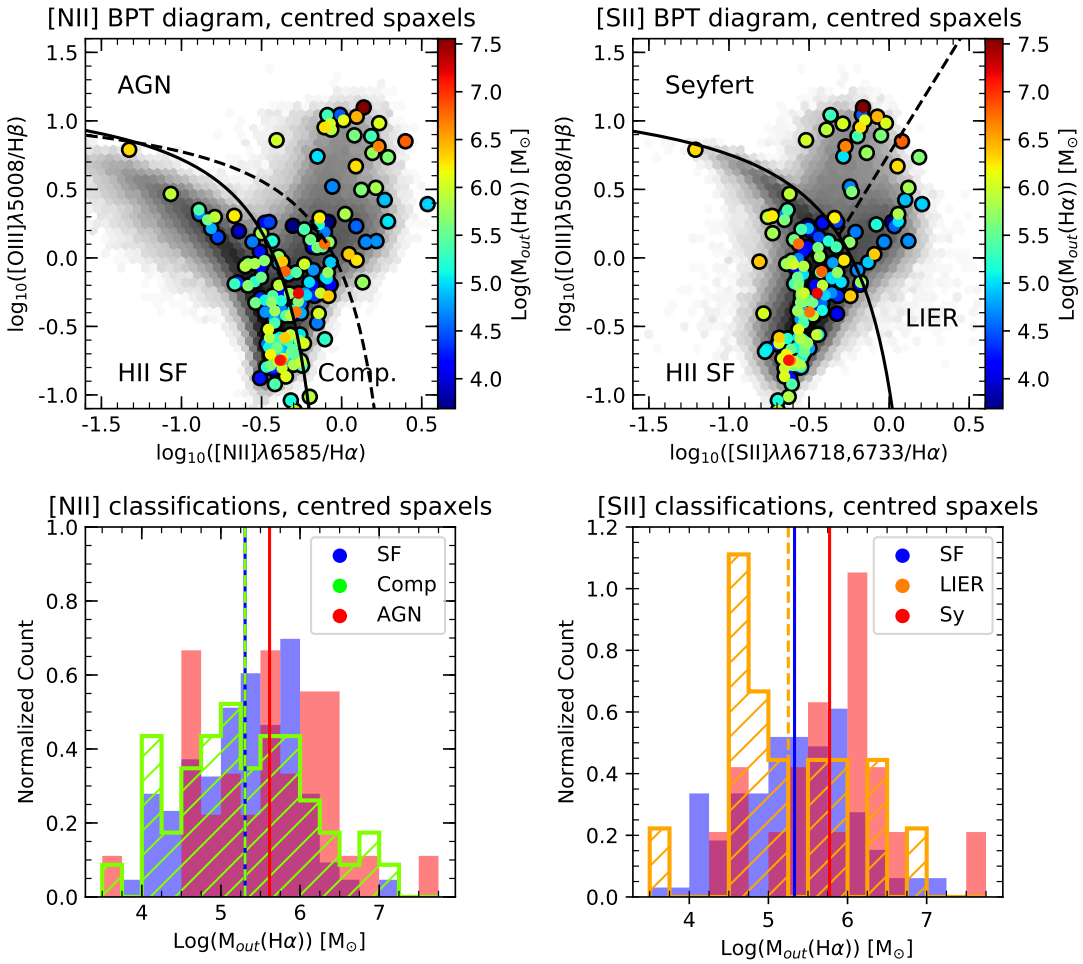


Figure 6.21: Top panels: Distribution of outflow driving sources on the [NII]- and [SII]-BPT diagrams (left and right panels, respectively), coloured by their outflow mass inferred from Hα, M_{outfl}(Hα). Bottom panels: Distributions of M_{outfl}(Hα), as split by their [NII]- and [SII]-BPT classified driving sources (left and right panels, respectively). We see here that M_{outfl}(Hα) is greater in AGN/Sy-driven outflows than SF-driven outflows by a factor ~2, suggesting AGN can more efficiently eject gas from a galaxy.

in part to SF, though a number may be AGN with large amounts of ongoing surrounding SF. The average M_{outfl}([OIII]) for the AGN classified driving sources is a factor of ~10 larger than those of the SF or Composite sources, suggesting [OIII] traces outflow mass in AGN-driven outflows more efficiently than in SF-driven outflows.

The large outflow masses inferred from [OIII] for AGN-driven outflows is not surprising, given that the hard radiation field of AGN increases the fraction of [OIII] relative to lower ionisation transitions. Whilst this result highlights the effectiveness of using [OIII] to trace the outflow mass in AGN-driven outflow, it also serves to highlight biases that might affect outflow

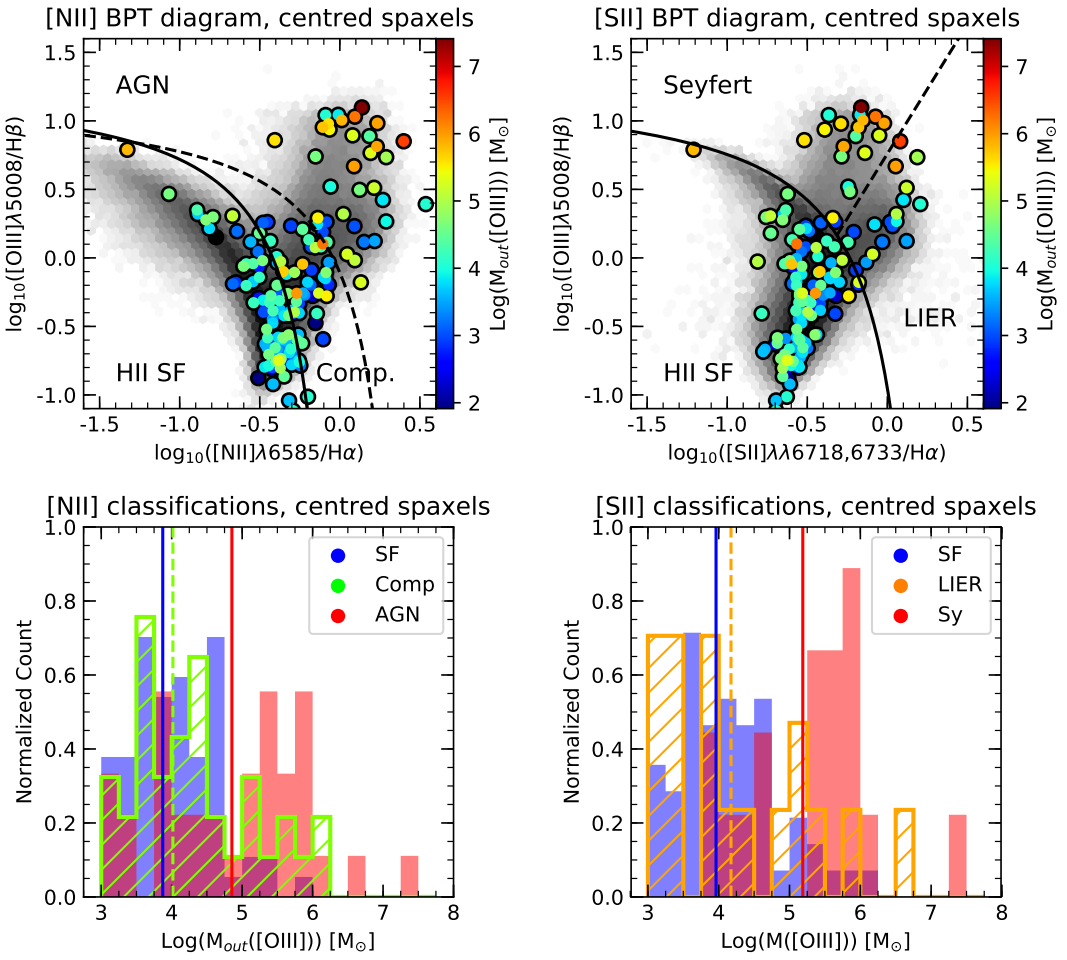


Figure 6.22: Top panels: Distribution of outflow driving sources on the [NII]- and [SII]-BPT diagrams (left and right panels, respectively), coloured by their outflow mass inferred from [OIII], M_{outfl}([OIII]]). Bottom panels: Distributions of M_{outfl}([OIII]), as split by their [NII]- and [SII]-BPT classified driving sources (left and right panels, respectively). We see here that M_{outfl}([OIII]) is significantly greater in AGN/Sy-driven outflows than SF-driven outflows by a factor ∼10, suggesting [OIII] more efficiently traces AGN-driven outflows than SF-driven outflows. This is unsurprising as the radiation field of AGN increase the fraction of [OIII] compared to lower ionisation lines.

surveys solely based upon the presence and analysis of the [OIII] profile.

In the [SII] distribution we see a similar trend, with those outflow driving sources classified as Sy dominating the high-mass end of the distribution. The M_{outfl}([OIII]) for driving sources classified as LIER show a large spread again in the distribution, but shows a similar peak to those driven by SF (the orange dashed and blue solid lines, respectively). Again, the Sy-driven outflows show M_{outfl}([OIII]) a factor of ∼10 larger than those of the other classifications,

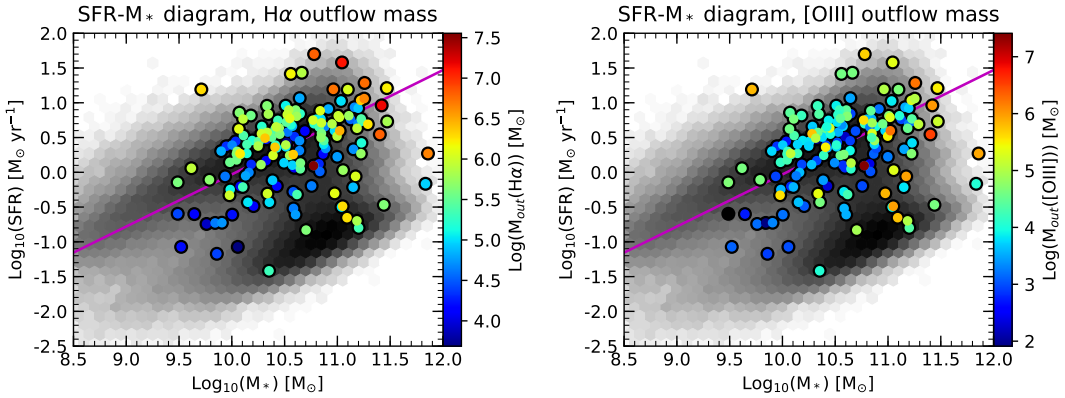
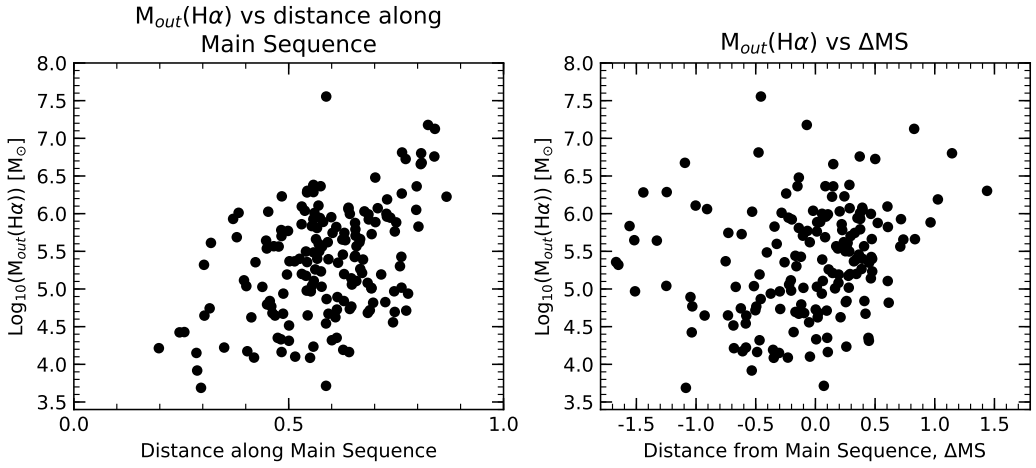


Figure 6.23: Distributions of the outflow host galaxies on the SFR-M_* plane, colour-coded by their Hα (left panel) and [OIII] (right panel) inferred outflow masses, M_{outfl}(Hα) and M_{outfl}([OIII]) respectively. For M_{outfl}(Hα), we find the highest M_{outfl}(Hα) outflows at the high-mass end of the MS and in galaxies in the starburst region, though moderate M_{outfl}(Hα) values are seen in galaxies throughout the diagram. For M_{outfl}([OIII]), we find the highest M_{outfl}([OIII]) outflows in the quiescent region, for driving sources previously classified as AGN, and in galaxies far above the MS in the starburst region. These results show that M_{outfl}(Hα) may better trace the outflow mass in outflows, with M_{outfl}([OIII]) better tracing the outflow mass in AGN-driven outflows as opposed to SF-driven outflows, though this is again unsurprising due to the hard radiation field of the AGN boosting [OIII] emission.

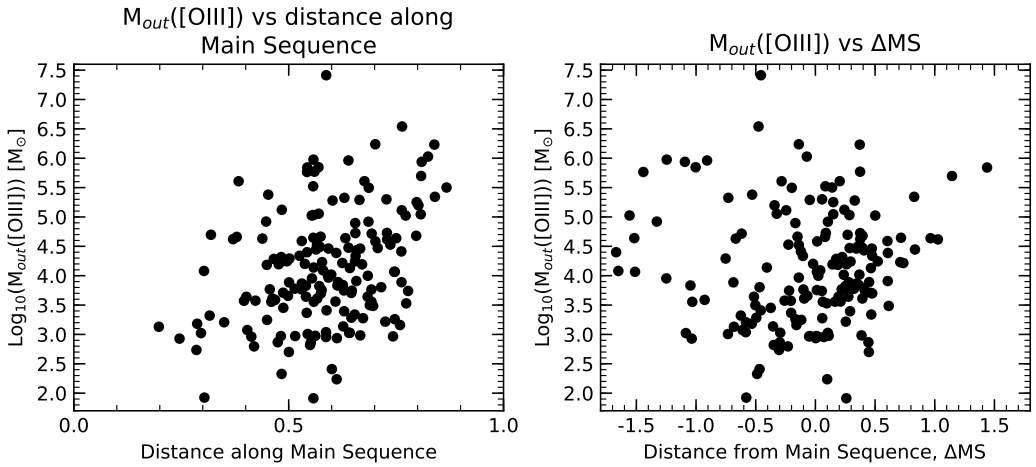
supporting the postulate that Sy-driven outflow masses are more efficiently traced in [OIII] then those of SF-driven outflows.

Returning to figure 6.23, the right hand panel shows the SFR-M_* diagram, with our outflow host galaxies colour-coded by their M_{outfl}([OIII]). As expected from the BPT diagrams, we see that the most massive outflows exist in galaxies in the green valley or on the red sequence, galaxies for which we have already ascertained an AGN to be the likely driving source of the outflow. We also again observe some of the massive outflows existing at the upper end of the MS, in similar locations to out most massive Hα outflows, indicating powerful driving sources.

In both panels of figure 6.23 we can see a tentative trend of outflow mass increasing with distance along the MS. To explore this, we plot in figure 6.24 both M_{outfl}(Hα) (top panels) and M_{outfl}([OIII]) (bottom panels) as functions of the distance along the MS (left panels) and ΔMS (right panels). Despite the moderate scatter, there does appear to be a trend of increasing outflow mass with distance along the MS for both M_{outfl}(Hα) and M_{outfl}([OIII]), though the trend is slightly tighter for M_{outfl}(Hα). As a function of ΔMS, however, we again see no real trends arising. These results again suggest that the properties of outflows are influenced by increasing values of both SFR and stellar mass, similar to the trends observed for R_{outfl} and



(a)



(b)

Figure 6.24: The distributions of the outflow mass inferred from Hα (top panels) and [OIII] (bottom panels) as functions of the distance along the MS (left panels) and ΔMS (right panels). Whilst we see no strong trends in the distributions of both outflow mass estimates as a function of ΔMS, we do see some tentative suggestion that more massive outflows are located in more massive and SF galaxies from the distribution of outflow masses as a function of distance along the MS, though there is a large amount of scatter.

τ_{dyn}.

6.6.2 Outflow mass rates

6.6.2.1 Outflow mass rate equations

To calculate the mass outflow rates for our outflows, we return to the basic principles set forth in section 1.3.2, namely equation 1.3.

Whilst we employ the same techniques as Carniani et al. (2015) for our calculations of our outflow masses, we here deviate from the techniques used by them as well as other authors (Cicone et al. 2015) and those employed in chapter 5 to determine mass outflow rates. We make this change as advances within the field have lead to the development of better techniques for the determination of outflow rates. Previously, the global outflow rate was calculated by summing the contribution each spaxel makes to the total outflow rate, with each spaxel assumed to be providing a contribution to the outflow rate. However, the basic tenet of this methodology (that summing contributions to the outflow rate from each spaxel can provide a robust estimate for the global outflow rate) will lead to an overestimation of the outflow rate, as one cannot simply add rates together to determine a single global rate. As such, we consider instead two methodologies that are likely to better explore the global mass outflow rate on spatially resolved scales

6.6.2.1.1 Radially expelled shells of gas

Inspired by Venturi et al. (2018), our first approach towards outflow rate calculations instead considers radially expanding shells of expelled gas. If one considers the basic principles of how an outflow could be driven from the plane of a galaxy, and thus how the expelled gas would expand, the simplest explanation is that of spherical shells, of some infinitesimal width dR, each containing some fraction of the total outflow mass dM. If we assume this gas does not experience significant deceleration from collision with either the ISM or the IGM or (though obviously not physically possible) due to the gravitational bounding of the galaxy, nor does it experience significant acceleration from radiation pressure or ram pressure from an entraining wind, one could compute a value for the outflow rate at some time in the past, τ_0, that the mass in the shell was emitted. The mass outflow rate associated with such a shell, initially found at radius R and expanding through a spherical surface at the same radius R, subtended by a solid angle Ω, such that the shell would be subsequently found at radius $R + dR$ (the width of the shell), can be computed as $\dot{M} = \Omega \mathrm{R}^2 \rho(\mathrm{R}) v(\mathrm{R})$. As such, if we can distinguish individual shells of width dR, we could explore the outflow rate history of our driving source. In practice, however, the shells themselves would be slowed through interactions, and due to projection effects and our observations of each galaxy being the 2D projection onto the plane of the detectors of the 3D outflow, we cannot resolve individual shells.

Despite these issues, we can recover an estimate for the outflow rate by instead considering annuli of width dR in our 2D plane, representative of shells of width dR. Whilst our annuli will never fully represent the mass contained only within the shell (they represent either an under-representation for the outer annuli or an over-representation for the inner annuli), the

average value of the outflow rates in each of the annuli should account for this issue. For each annuli, between radii R and $R + dR$, we calculate the outflow rate as

$$\dot{M}(R < r \leq R + dR) = \sum_i M_i \frac{v_{max}}{dR} \tag{6.3}$$

where M_i is the mass contained in spaxel i that falls within the annulus, and v_{max} is the maximum velocity of all contained spaxels, defined similarly as before as $|v|$ + FWHM/2. The majority of outflows are never driven at velocities that exceed the escape velocity of the host galaxy, and as such will be decelerated and rain back down onto the galaxy disk. Hence, in annuli that trace this gas, the calculated outflow rates will be massively underestimated, due to the reliance on the outflow velocity. To account for this, we must determine a cutoff radius, external to which we do not consider the outflow rate representative of the initial outflow rate when the shell was emitted. For this purpose, we we use the outflow extents we previously determined in section 6.5. The outflow rates determined in this way will be those out of the vicinity of the galactic disk or out of the vicinity of the galaxy, and as such will be much lower than the effective outflow rate (i.e. the rate at which the mass of gas entrained within the outflow permanently leaves the galaxy).

6.6.2.1.2 Integrated mass outflow rate

Equally, one can compute the mass outflow rate by considering the total mass in the outflow, $M_{outfl,\,T}$, and estimating the dynamical timescale, τ_{dyn}, from the total extent of the outflow, R_{outfl}, and the maximum outflow velocity exhibited by the outflow spaxels, v_{outfl}. In fact, we have already defined τ_{dyn} from R_{outfl} and v_{outfl} in section 6.5, and can recover $M_{outfl,\,T}$ from our spatially resolved outflow mass maps by simply summing over all spaxels contained within the outflow. We then substitute these values into equation 1.3 to recover our outflow rates. This methodology is similar to the methods more typically used in other works, where spatially resolved information is not available.

6.6.2.2 Computing and comparing the outflow mass rates

The bottom panels in figure 6.20 present graphical representations of the radial bin methodology, showing the outflow rate in each bin of width dR=0.5 arcsec, between radii R and R+dR, using equation 6.3. As discussed above, we see the expected trend of outflow rates decreasing with radius, due to the inner annuli containing mass from not just the radially expanding shells in these radial bins, but also a partial contribution from previously emitted shells. To consequently determine the time averaged global mass outflow rates, $\dot{M}_{Outfl,\,Avg.}(\mathrm{H}\alpha)$ and $\dot{M}_{Outfl,\,Avg.}([\mathrm{OIII}])$, we consider the average mass outflow rate of all bins internal to the

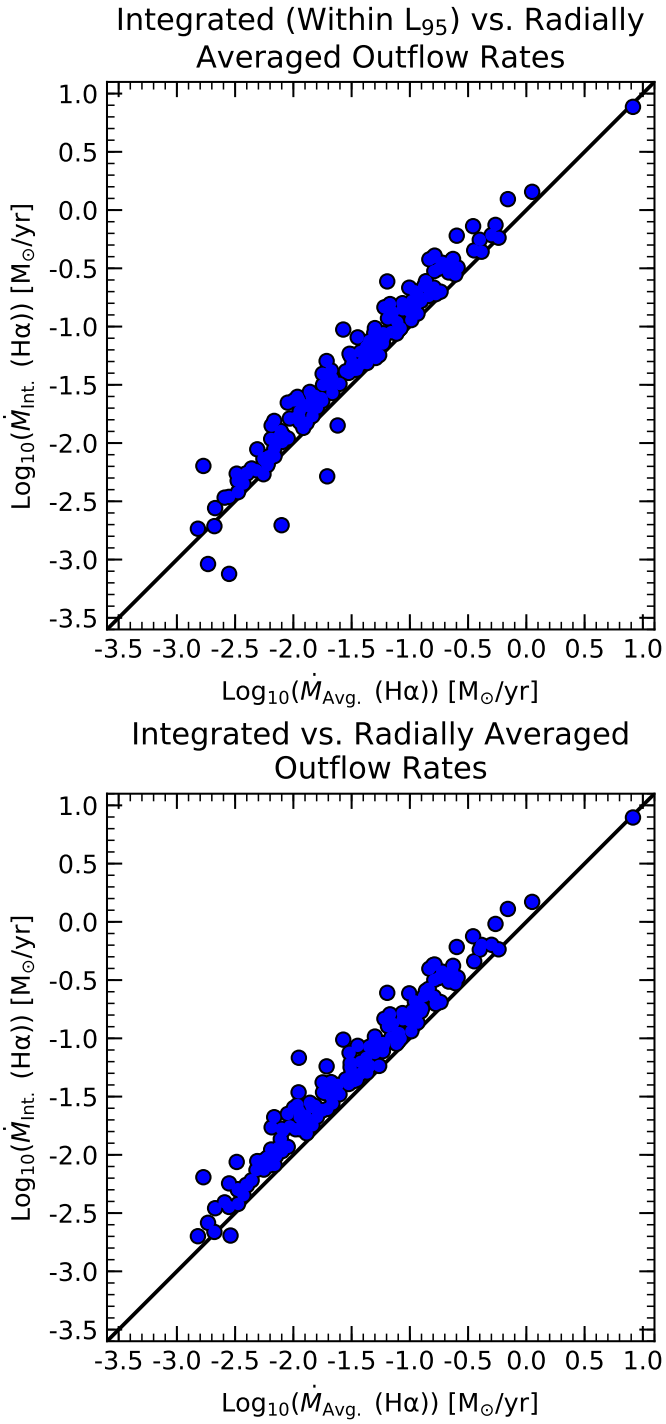


Figure 6.25: Outflow mass rates, as inferred from Hα (top panel) and [OIII] (bottom panel), determined using the two different methodologies discussed in section 6.6.2.1, namely the radially binned average and integrated outflow mass rate estimations. The computed outflow mass rates for each outflow are plotted as blue circles. The black line represents a 1:1 ratio. We can see that the outflow mass rates of the two methodologies are in close agreement, suggesting the new radial bin average methodology is able to recover good estimates of the outflow mass, and that these outflow rates are likely to agree with those in literature recovered for data where spatially resolved information is unavailable.

outflow extent radius. In the bottom panels of figure 6.20 the cutoff radius is shown as the vertical dashed black line, whilst the average outflow mass rate is shown as the horizontal dot-dashed black line. The integrated outflow mass rates, $\dot{M}_{Outfl,\ Int.}(\mathrm{H}\alpha)$ and $\dot{M}_{Outfl,\ Int.}([\mathrm{OIII}])$, are simply calculated from the total gas mass entrained within the outflows.

We compare these two outflow rate calculations in figure 6.25 for both Hα and [OIII], where the solid black line represents a 1:1 ratio. As can be seen, our outflow rates show strong agreement, with the integrated outflow rates providing slightly higher estimates for the outflow rates. This suggests that, when spatially resolved data is available, calculations of outflow mass rates from considerations of radial bins will provide comparable results to those of previous studies that lacked spatially resolved data. From this point on, we use the outflow rates calculated using our radial bin methodology.

6.6.2.3 Distribution of outflow mass rates on the BPT-diagrams and SFR-M$_*$ plane

Having calculated the outflow mass rates, we now explore the distribution of the Hα and [OIII] inferred outflow mass rates, $\dot{M}_{outfl}(\mathrm{H}\alpha)$ and $\dot{M}_{outfl}([\mathrm{OIIII}])$, as a function of position on the BPT and SFR-M$_*$ diagrams.

Figure 6.26 shows the locations of the outflow driving sources on the [NII]- and [SII]-BPT diagrams, coloured by each outflow's $\dot{M}_{outfl}(\mathrm{H}\alpha)$. Starting with the [NII]-BPT classifications, we observe that the driving sources of those outflows with higher outflow rates tend to be classified as AGN. This can be seen in both the [NII]-BPT diagram, as well as the outflow rate distribution histogram (red filled histogram). Similarly to the outflow masses, we see that the SF and Composite classified driving sources show similar distributions (blue filled and green hashed histograms, respectively), and that the average outflow rates of the SF and Composite classified sources are very similar (blue solid line and green dashed line, respectively). Again, this is suggestive that for a vast majority of driving sources classified as Composite, the outflows are likely SF-driven as opposed to AGN-driven.

A similar set of trends is shown in the [SII]-BPT classifications, with Sy-driven outflows exhibiting greater mass outflow rates (red filled histogram) when compared to the distributions of the SF and LIER classified driving sources (blue filled histogram and green hashed histograms, respectively). Again, similarly to the observed trend in the $M_{outfl}(\mathrm{H}\alpha)$ distributions, the SF and LIER classifications show very similar average mass outflow rates. In both cases, the AGN-driven outflow rates seem to be around a factor of ~2-3 greater (0.3-0.5 dex logarithmically) greater than those of the SF or Composite/LIER classifications. The finding that AGN boost the outflow rate by a significant factor, relative to SF, is in agreement with and reinforces similar findings obtained for the molecular phase of outflows (Cicone et al. 2014; Fluetsch et al. 2019).

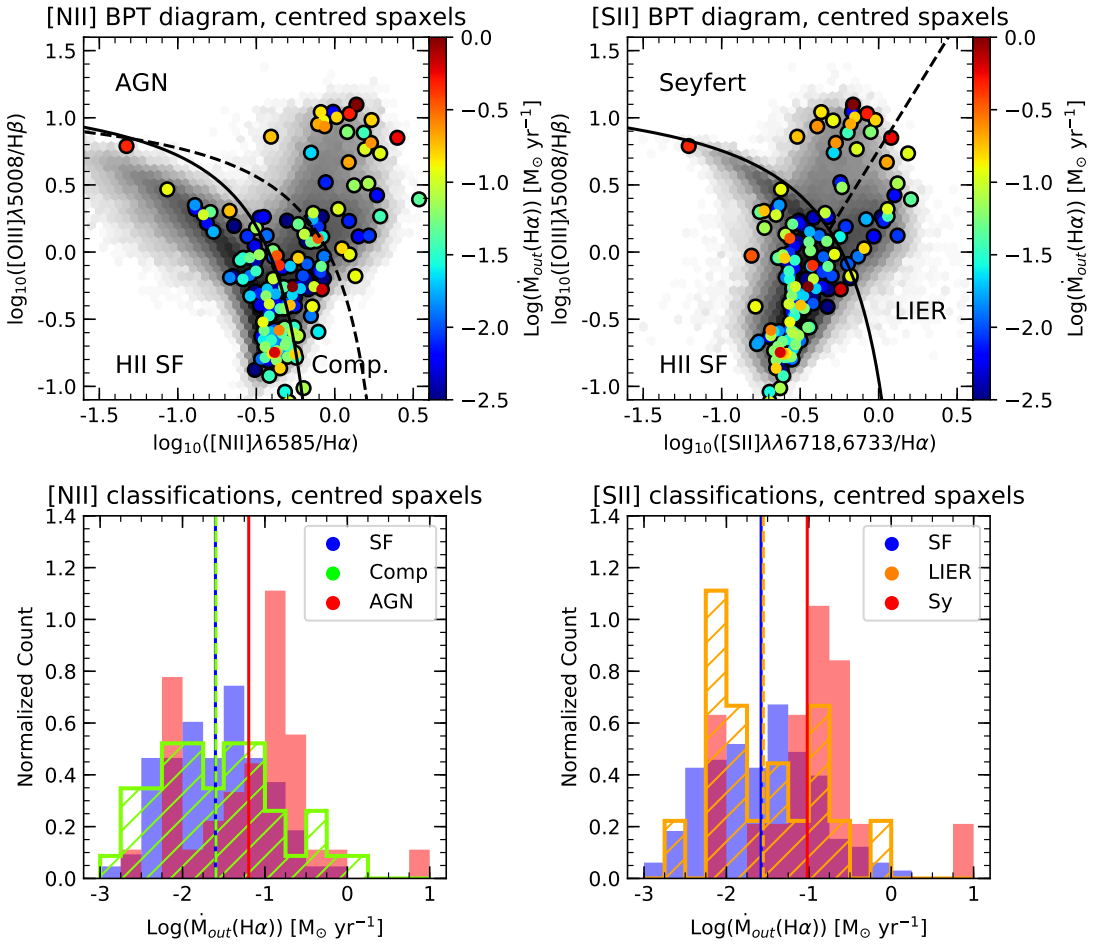


Figure 6.26: Top panels: Distribution of outflow driving sources on the [NII]- and [SII]-BPT diagrams (left and right panels, respectively), coloured by their outflow mass rates inferred from Hα, $\dot{M}_{outfl}$(Hα). Bottom panels: Distributions of $\dot{M}_{outfl}$(Hα), as split by their [NII]- and [SII]-BPT classified driving sources (left and right panels, respectively). We see here that, similar to the distributions of the outflow mass inferred from Hα, $\dot{M}_{outfl}$(Hα) is, on average, greater in AGN/Sy-driven outflows than SF-driven outflows by a factor ~2. The Comp. and LIER classified driving sources show a similar distribution and average to the SF-driven outflows, suggesting many of these may in fact be SF-driven outflows.

Finally for $\dot{M}_{outfl}$(Hα) we consider the locations of the outflow hosts on the SFR-M$_*$ diagram, as shown in figure 6.28. In this diagram, it appears that outflows with modest $\dot{M}_{outfl}$(Hα) are found throughout. Those outflows with the highest outflow rates are again found at the upper end of the MS and in the starburst region above it, but outflows all along the MS and in the red sequence also exhibit significant (with respect to the local Universe) outflow rates.

Figure 6.27 shows both BPT diagrams, colour-coded by the $\dot{M}_{outfl}$([OIII]) of each galaxy,

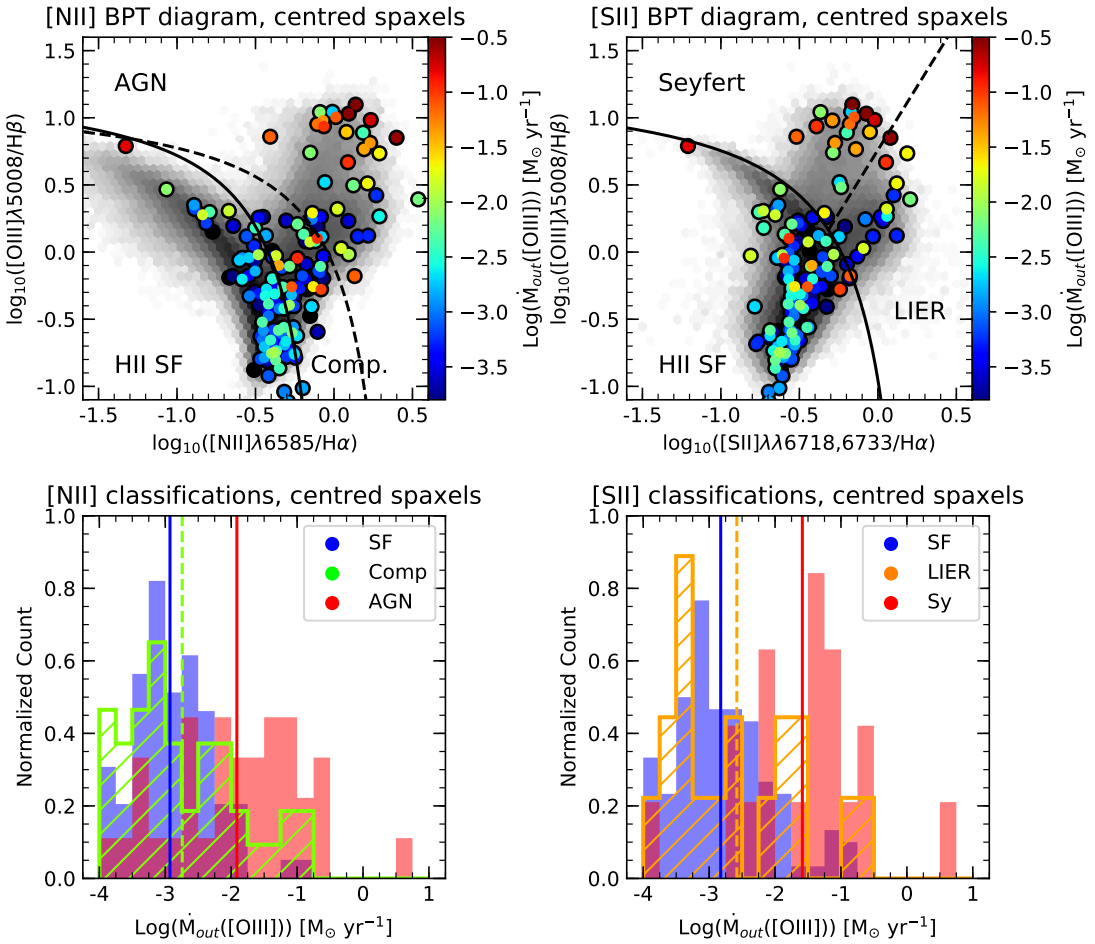


Figure 6.27: Top panels: Distribution of outflow driving sources on the [NII]- and [SII]-BPT diagrams (left and right panels, respectively), coloured by their outflow mass rates inferred from [OIII], $\dot{M}_{outfl}$([OIII]). Bottom panels: Distributions of $\dot{M}_{outfl}$([OIII]), as split by their [NII]- and [SII]-BPT classified driving sources (left and right panels, respectively). We see here that, similar to the distributions of the outflow mass inferred from [OIII], $\dot{M}_{outfl}$([OIII]) is, on average, significantly greater in AGN/Sy-driven outflows than SF-driven outflows by a factor ∼10. This further supports the postulate that [OIII] is a better tracer of outflow mass in AGN-driven outflows, likely due to the AGN radiation field boosting the [OIII] emission.

and the distributions of the $\dot{M}_{outfl}$([OIII]) split by the driving source BPT classification. Starting with the [NII]-BPT diagram, we can see that almost all galaxies with significant $\dot{M}_{outfl}$([OIII]) are classified as AGN, further reinforcing the results from the Hα-inferred outflow mass rates and supporting the postulate that an AGN driving source significantly boosts the outflow rate. Once again, it is not surprising that this effect is seen more strongly in [OIII] than in Hα, as AGN radiation tends to enhance the [OIII] emission. Considering the distributions show in the bottom panels, there is a clear dichotomy between the $\dot{M}_{outfl}$([OIII]) of SF-driven and AGN-driven

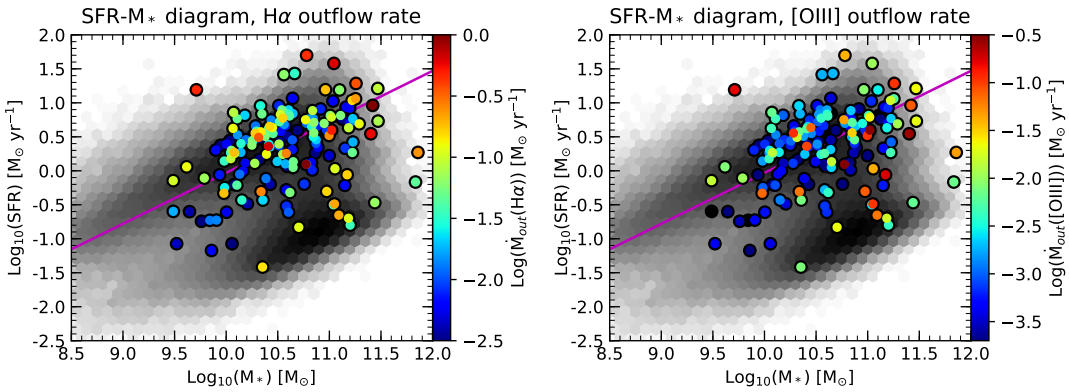
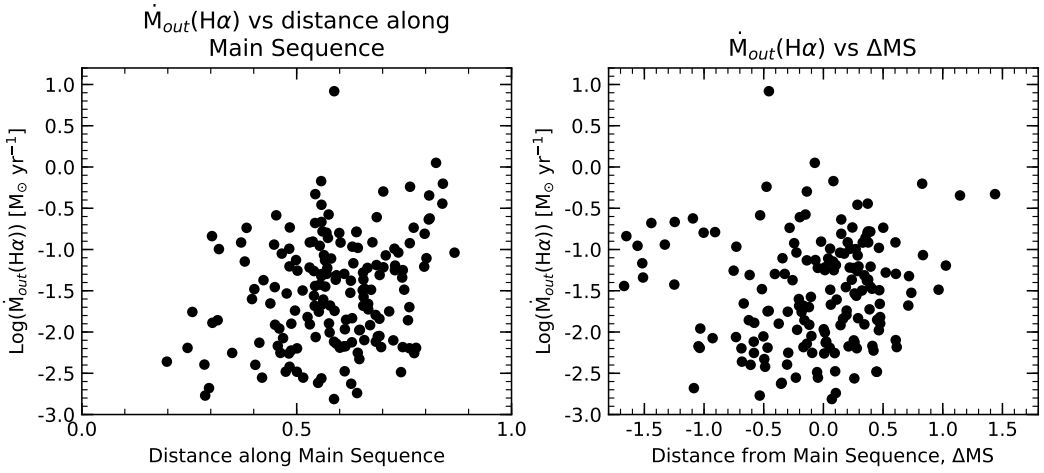


Figure 6.28: Distributions of the outflow host galaxies on the SFR-M$_*$ plane, colour-coded by their Hα (left panel) and [OIII] (right panel) inferred outflow mass rates, $\dot{M}_{outfl}$(Hα) and $\dot{M}_{outfl}$([OIII]) respectively. For $\dot{M}_{outfl}$(Hα), we find the highest $\dot{M}_{outfl}$(Hα) outflows at the high-mass end of the MS and in galaxies in the starburst region, though moderate $\dot{M}_{outfl}$(Hα) values are also seen in galaxies in the quiescent region, associated with AGN-driven outflows. For $\dot{M}_{outfl}$([OIII]), we find the highest M_{outfl}([OIII]) outflows in the quiescent region, for driving sources previously classified as AGN, in galaxies at the high-mass end of the MS and in galaxies far above the MS in the starburst region. These results show that $\dot{M}_{outfl}$(Hα) may better trace the outflow mass rate in outflows in general, with $\dot{M}_{outfl}$([OIII]) better tracing the outflow mass rate in AGN-driven outflows as opposed to SF-driven outflows, though this is again unsurprising due to the hard radiation field of the AGN boosting [OIII] emission.

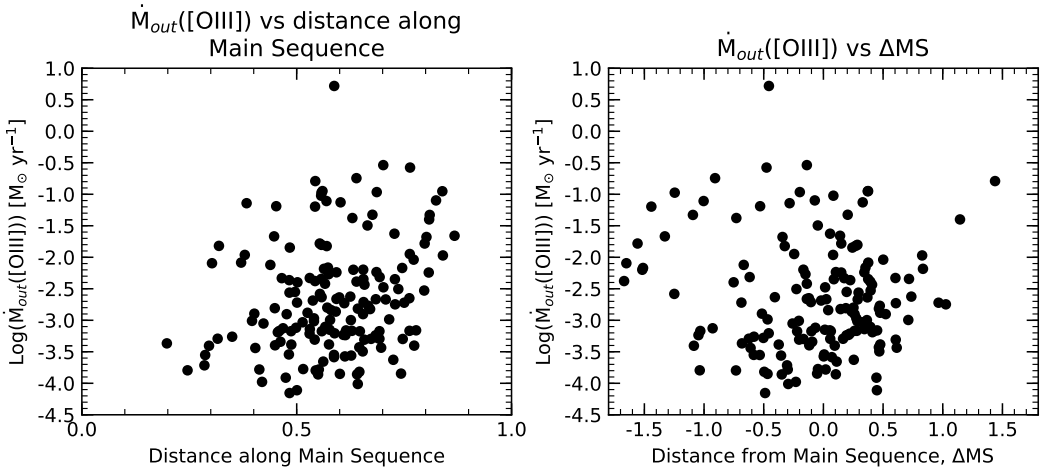
outflows (blue filled and red filled histograms, respectively), with the distribution of outflow rates of Composite classified driving sources (green hashed histogram) again appearing most similar to that of SF classified sources, whilst the AGN-driven outflow rates are now higher by nearly an order of magnitude.

In the [SII]-BPT diagram, this trend continues and is further emphasised. The driving sources of the outflows with the greatest $\dot{M}_{outfl}$([OIII]) are located in the Sy locus on the BPT diagram, with two outflows with high $\dot{M}_{outfl}$([OIII]) having their driving sources classified as LIER but close to the Sy demarcation line. Considering the distributions, the dichotomy of the [NII]-BPT classifications is again present, with $\dot{M}_{outfl}$([OIII]) of Sy-driven outflows dominating the upper end of the distribution (red filled histogram). The outflows associated with LIER classified sources show similar $\dot{M}_{outfl}$([OIII]) to those of the SF classified sources, with their average $\dot{M}_{outfl}$([OIII]) being very similar, once again (orange dashed line and blue solid line respectively). Finally, the average mass outflow rates for Sy classified sources are a factor $\sim$10 greater than those of outflows with SF or Composite/LIER sources, similar to AGN-driven outflows in the [NII] distributions.

Returning to figure 6.28, we consider the locations of the outflow host galaxies on the



(a)



(b)

Figure 6.29: The distributions of the outflow mass rates inferred from Hα (top panels) and [OIII] (bottom panels) as functions of the distance along the MS (left panels) and ΔMS (right panels). Here, we see no strong trends in any of the figures, suggesting that the outflow mass rate is a quantity influenced by the outflow driving source alone, and not necessarily by the properties of the host galaxy.

SFR-M$_*$ plane as a function of their $\dot{M}_{outfl}$([OIII]). Here, we can see that outflows exhibiting low $\dot{M}_{outfl}$([OIII]) typically exist along the MS, supporting previous results that [OIII] is a poor tracer of SF-driven outflows. In contrast, some of the largest $\dot{M}_{outfl}$([OIII]) are found in outflows that reside on the red sequence, for which we have already determined an AGN driving source, again supporting the previous results that [OIII] preferentially traces AGN-driven outflows due to the AGN radiation enhancing [OIII] emission. We do observe, however, a number of outflows with high $\dot{M}_{outfl}$([OIII]) in both starburst galaxies above the MS and

in galaxies at the high-mass end of the MS. These galaxies also exhibit high $\dot{M}_{outfl}(\mathrm{H}\alpha)$, suggesting that the outflow driving sources in these galaxies must be powerful.

In contrast to a number of the other outflow properties we have explored, the outflow mass rates don't seem to show any significant trend with either location along the MS or ΔMS. This is explored further in figure 6.29, where we plot $\dot{M}_{outfl}(\mathrm{H}\alpha)$ and $\dot{M}_{outfl}([\mathrm{OIII}])$ (top and bottom panels, respectively) as functions of the distance along the MS and ΔMS (left and right panels, respectively). In all of the plots, no significant trends arise, suggesting that the outflow mass rates are likely not influenced by the properties of their host galaxies.

6.6.3 Outflow mass loading factor

To consider the effect our ionised outflows have on the SF activity in their host galaxy, we utilise the outflow mass loading factor, a crucial parameter in models of galaxy evolution. This quantity compares the amount of gas being ejected by the outflow, $\dot{M}$, to the SFR of the gas in the ISM, with the mass loading factor being given by the ratio between these two quantities, $\eta = \dot{M}/\mathrm{SFR}$. As we have outflow mass rates based on both Hα and [OIII], we calculate the mass loading factor for both. However, we must recall that this outflow rate only refers to the ionised phase; the neutral atomic and molecular phases can add a major contribution to the outflow mass rate.

The top panels of figure 6.30 show histograms of the mass loading factor for each of our outflows, determined from their Hα inferred mass outflow rates, and colour coded by the [NII]- and [SII]-BPT classifications (left and right panels, respectively) of their outflow driving sources. As can be seen, only a few AGN-driven outflows show mass loading factors greater than 1, suggesting their outflows (or the AGN driving them) are seriously disrupting the SF in their host. For the remaining AGN-driven outflows, the leading factor falls between $0.1 < \eta < 1$, suggesting that although they do not have a significant affect on the ongoing SF of their hosts, they are playing a role in quenching. A few of the driving sources classified as Composite or LIER also fall into this range, suggesting the driving source may be an AGN. For the vast majority of the Composite classified sources, as well as the SF Classified sources, the outflows show loading factors of around, on average, 0.01. This suggests that for SF-driven outflows, the outflows themselves are not playing any significant part in quenching the ongoing SF in their host.

The bottom panels of figure 6.30 instead shows the mass loading factors for our outflows determined from their [OIII] inferred mass outflow rates, again coloured by their [NII]- and [SII]-BPT classifications (left and right panels, respectively). Whilst we observe similar trends to that of the mass loading factors determined using the Hα inferred mass outflow rates, overall the mass loading factors are much smaller. In fact, we only observe one galaxy (hosting an

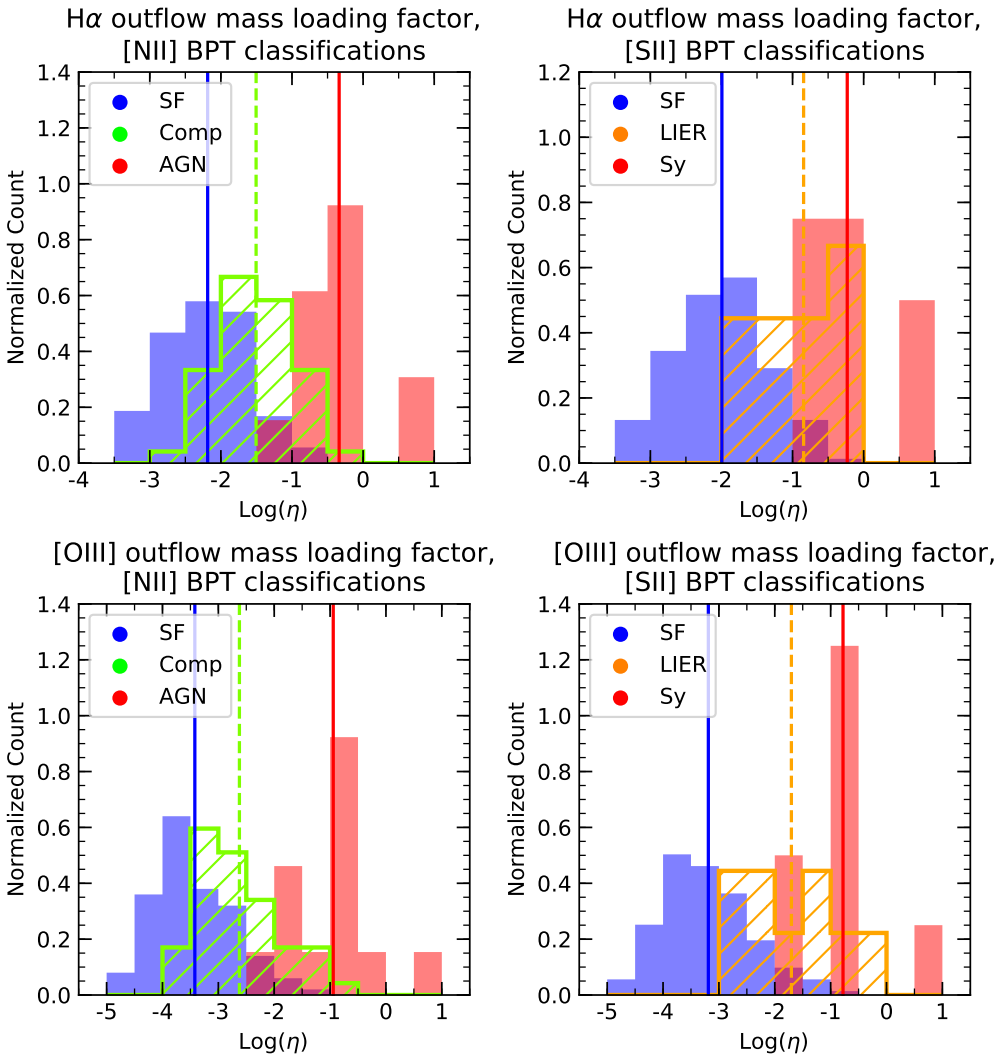


Figure 6.30: Histograms of the outflow mass loading factor determined from the Hα (top panels) and [OIII] (bottom panels) inferred outflow mass rates, split by the outflow driving source determined using the [NII]- and [SII]-BPT diagnostics (left and right panels, respectively). We can clearly see that AGN/Sy-driven outflows have a much larger effect in quenching the SF of their host, either through the outflow itself or the AGN presence. SF-driven outflows seem to have very little effect in the quenching the SF of their host. Comp. and LIER classified driving sources seem to straddle the two distributions, with intermediate outflow mass loading factors, suggesting the driving sources of these outflows may be stronger SF-driven or weaker AGN-driven outflows. One caveat, as discussed below, is that these outflow mass loading factors were only calculated using the ionised component of the outflow, which may account for only a small fraction of the total outflow.

AGN-driven outflow) with a mass loading factor above 1, with most AGN-driven outflows falling in the range $0.01 < \eta < 0.2$, suggestive of an even lower effect of the SF of their host. The Composite and LIER classified driving sources are again in an intermediary location, with SF driving sources showing little to no impact on their host, with average loading factors around 0.001.

However, as discussed above, one caveat here is that we are using the ionised mass outflow rates as a proxy for the global mass outflow rate of the galaxy, as we currently do not have information on the other phases of the outflow. Many previous studies of both the ionised and molecular phases of outflows have shown that the ionised phase may only account for a small fraction of the total mass entrained within an outflow (Carniani et al. 2015; Fiore et al. 2017; Fluetsch et al. 2019). Thus if we could include these additional phases, the mass loading factors would likely be much higher than we observe here.

6.7 Conclusions

In this section, we have located and characterised ionised outflows driven from nuclear regions of ∼4600 local galaxies from the MaNGA survey. The search for outflows resulted in the detection and classification of 168 ionised outflows in galaxies spanning the stellar mass range $\sim 10^{9.5}-10^{12}$ $M_\odot$ and the SFR range ∼0.03−100 $M_\odot$ yr^{-1}, traced by broad emission lines in Hα and/or [OIII]λ5008. We also trace the [SII] doublet in the outflow to determine the density of the gas entrained within the outflow. For all of the host galaxies, we trace all of the required nebular lines to characterise them through BPT diagnostic in the narrow (disk) component, including the [SII] doublet again enabling us to study the gas density.

We have obtained the following results:

(i) The median BPT and WHAN diagnostic classifications of all of the spaxels that constitute the narrow (disk) component for each outflow host galaxy suggest that our outflow host galaxies form a representative sample of galaxies in the local Universe, when compared to the results of SDSS DR7. This suggests that outflows are prevalent in all types of galaxy in the local Universe.

(ii) The median BPT and WHAN diagnostic classifications of the narrow component spaxels that fall within the central PSF of the host galaxy, that likely trace the outflow driving sources, indicate the occurrence of AGN in outflow host galaxies is twice that in a normal sample of local galaxies. This suggests, unsurprisingly, that many outflows are AGN-driven. However, a large fraction of our outflows are classified as SF-driven, indicating the propensity of SF-driven outflows in the local Universe.

(iii) In the SFR-M_* plane, we find our outflow host galaxies are most commonly found on or above the MS, though a small number do fall within the green valley and quiescent regions. When considering the ΔMS distribution, we see the outflow hosts peak above the MS in the starburst region. When splitting our galaxies by their 'driving-source' BPT classifications, we see that SF-driven outflow hosts continue to peak above the MS in the starburst region, whilst AGN-driven outflows tend to be located below the MS, but more widespread in their distribution with some on or slightly above the MS. These results suggest that SF-driven outflows are more common in the presence of a starburst, and that AGN are an important driving mechanism for outflows but are more effective in driving outflows in galaxies on the MS, as more ISM gas is available to couple to the AGN winds and radiation pressure.

(iv) The total average densities of gas entrained within outflows, n_e=295 cm^{-3}, is denser than gas in the disk, n_e=98 cm^{-3}, by a factor of ∼3. The density of the gas in the disk does not seem to be affected by outflow presence. When split by driving source, the broad components show the gas in SF-driven outflows is, on average, slightly less dense than the gas in AGN-driven outflows. When considering radial trends, the density of the gas in both the outflow and the disk decrease with radius. Again splitting the broad component by driving source we see that the density in SF-driven outflows is lower than that of AGN-driven outflows at all radii.

(v) Outflow velocities in SF-driven outflows generally do not exceed a few hundred km s^{-1}, whilst AGN driven outflows can exceed 1000 km s^{-1}. The velocity of outflows show little to no trend with distance along the MS, though they do show a tentative trend of increasing with ΔMS.

(vi) Outflow extents show that, on average, AGN-drive outflows are slightly more extended, but that there are also a few very extended SF-driven outflows, likely due to the driving source being distributed on larger scales. Whilst no obvious trend is seen as a function of ΔMS, the extent seems to show a significant trend of increasing extent with distance along the MS. This suggests that outflow extent is dependent on both host stellar mass and star formation rate.

(vii) Outflow dynamical timescales are found to be only slightly longer for SF-driven outflows than AGN-driven outflows, likely due to the flickering nature of AGN. Similar to the outflow extent, the timescale shows a strong correlation with distance along the MS, suggesting the timescale of outflows increases as a function of both host stellar mass and SFR. No significant trend is seen as a function of ΔMS.

(viii) Outflow masses, as inferred from Hα, tend to be higher in AGN-driven outflows than SF-driven outflows by a factor of ∼2. However, outflow masses inferred from [OIII] show

AGN-driven outflows are more massive by a factor of ∼10. These results suggest that AGN-driven outflows are more massive, and that AGN-driven outflows are preferentially traced by [OIII], unsurprisingly as the radiation field from AGN increases [OIII] emission relative to lower ionisation lines.

(ix) Similarly to outflow masses, outflow rates inferred from Hα tend to be lower in SF-driven outflows than AGN-driven outflows by a factor ∼2−3, whilst outflow rates inferred from [OIII] are a higher in AGN-driven outflows than in SF-driven outflows by a factor of ∼10. Again, this suggests that AGN-driven outflows are preferentially traced by [OIII] due to the AGN boosting the [OIII] emission.

(x) The outflow mass loading factor of our outflows, when split by BPT, indicate that AGN-driven outflows have a greater impact on the SFR in their host than SF-driven outflows. Whilst our outflow mass loading factors tend to be low, we have only calculated them using the mass contained in the ionised component of our outflows. As such, the outflow mass loading factors are likely to be much greater than we recover. These results suggest that AGN-driven outflows play a larger role in the quenching SF in their host than SF-driven outflows.

Chapter 7

Conclusions

7.1 Summary

Throughout this thesis I have focused upon the role outflows play in driving galactic evolution. Specifically, I have focused upon a new channel of SF, where SF occurs *within* galactic outflows (chapters 4 and 5), before exploring the general properties of outflows and their host galaxies, and the effect of (both positive and negative) outflow feedback on the outflow host galaxy. In this section, I summarise the main findings of this thesis.

Can SF occur within gas entrained within an outflow, as predicted by recent theoretical models? In chapter 4 we explored a known massive outflow in the local galaxy IRAS F23128−5919 through a combination of spatially resolved MUSE data and long-slit, high resolution X-shooter spectroscopic data. Using these two datasets, we traced the ionised phase of the outflow through multiple component fits to the optical nebular emission lines.

We analysed the locations of the nebular emission lines on the [NII]-, [SII] and [OI]-BPT diagrams and found that the broad emission line components tracing the outflow are located in the SF loci. These diagrams alone are insufficient to unambiguously prove the existence of ongoing SF *within* the outflow alone, and so we employed NIR diagnostics, afforded to us by the X-shooter data, to further support the postulate of *in situ* SF. These diagnostic diagrams do indeed further confirm that the ionisation of the gas entrained within the outflow is due to SF as opposed to AGN/shock ionisation. Finally, we explored the ionisation parameter of the

outflowing gas, which again confirmed that the SF-like ionisation of the gas in the outflow *is* due to stars *within* the outflow, as opposed to stars in the disk of the galaxy.

The final piece of evidence confirming ongoing SF within the outflow was the detection of a young stellar population with the kinematical fingerprints of formation inside the outflow. These were obtained through spectral fitting of the optical X-shooter spectra, dominated by features tracing young stars such as the Balmer lines, CaII triplet and HeI absorption. The kinematics of these features, each tracing different types of stars, showed blue-shifted velocities with respect to the stellar population of the disk, confirming their formation within the outflow.

How prevalent is this new mode of SF in outflows? In chapter 5 we followed up on the work of chapter 4 by exploring outflows in an unbiased sample of galaxies drawn from the MaNGA survey. We searched for centrally driven outflows in all galaxies in the sample, finding a sub-sample of 37 galaxies which possessed sufficient SNR in all of the nebular emission lines of the outflowing gas required to use BPT diagnostics.

Of the sub-sample, we found that around 30% of our outflows have median BPT classifications that fall into the SF loci on the [SII]- and [OI]-BPT diagrams, and in either the SF or Composite loci on the [NII]-BPT diagram, suggesting the outflowing gas is primarily excited by the UV radiations of hot, young stars. The spatially resolved BPT maps also reveal that, even in outflowing gas primarily excited by AGN-like ionisation, half of our outflows show at least some fraction (>10%) of spaxels indicating outflowing gas excited by SF.

Similarly to chapter 4, the BPT diagnostics alone are insufficient at conclusively proving that the excitation due to young stars is due to their presence within the outflowing gas, and may instead be due to young stars in the galactic disk. As such, we again employ the ionisation parameter diagram. Our results for this diagram show that the ionisation parameter of our outflowing gas is consistent with or slightly higher than that of SF galaxies in the SDSS DR7 dataset. As such, the excitation must arise from stars within the outflow.

To discount the possibility of shock ionisation possibly producing the SF-like line ratios we observe, we employ the MAPPINGS III model library of shocks, spanning a broad parameter space. We showed that shocks cannot produce the ionisation observed, both in terms of line ratios and in terms of ionising photon budget.

All of these results combined have shown that SF within galactic outflows is not as rare of a phenomenon as originally believed, with an occurrence rate of around 30% among outflows.

Why have detections of *in situ* SF remained elusive? In chapter 5 we considered the reasons for why this channel of SF has remained elusive until recently. One of the key issues with detecting

SF within an outflow lies in the weightings afforded by different excitation mechanisms to the BPT diagrams. We showed than, even in the presence of prominent SF within an outflow, the presence of even a weak AGN would dominate the BPT diagnostics. From simple energetics considerations, we showed that the Hα luminosity of an AGN would exceed the Hα luminosity of SF in the galaxy spheroid by a factor of around 5, effectively overwhelming SF in these diagrams. Furthermore, in the presence of an AGN, the [NII], [SII] and [OI] emission is more luminous than in SF regions, again skewing the diagrams towards an AGN classification.

In many outflow studies, the [NII]-BPT is the only diagnostic used to classify outflows. However, this diagram is sensitive to both the ionisation parameter and nitrogen abundance, leading to possible misclassifications of outflows as Composite rather than SF.

One additional consideration is that a large number of surveys designed to study outflows search for broad wings in [OIII] emission as signatures of outflows. As [OIII] emission is enhanced by the radiation fields of AGN, this line typically dominates in AGN-driven outflows, biasing studies to these specific types of outflows.

Finally, in the integrated, spatially unresolved (or poorly resolved) spectra typical of previous studies of outflows, the light weighting further enhances the bright, central AGN dominated emission relative to the weaker SF regions. This issue is compounded further for studies at high-z, as spatial resolutions are even lower than for local galaxies.

Despite these issues, potential indications of SF within outflows has been detected across the Universe, and other studies are beginning to expand upon our work, finding *in situ* SF in detailed studies of individual outflows (Venturi et al. *in prep.*; Fluetsch et al. *in prep.*) and in other studies of outflows in the MaNGA sample (Rodríguez del Pino et al, 2019).

What are the consequences of this new mode of SF for the evolution of galaxies, both in the local and high redshift Universe? In chapter 5 we considered the SFR of the gas entrained within the outflows compared to the SFR of the host galaxy, finding that SF within outflows could account for between 10% and 30% of the global SF of the galaxy.

To explore this further, in chapter 5 we investigated the connection between outflow mass rates and SF within the outflow. The data revealed a clear, linear correlation between the two quantities. The extrapolation of the observed relationship to the typical observed ionised outflow rates of high-z galaxies implies that large SFRs should be present in the outflows of these distant galaxies.

The stars that we expect to form within galactic outflows would possess entirely different dynamical properties to those formed in galactic discs, having primarily radial orbits.

Whilst outflows in the local Universe tend to be less powerful, and most stars are expected to be gravitationally bound, in powerful high-z outflows we would expect some of these stars

to posses sufficient velocity to escape the galaxy and enter the IGM. These stars could explain observations of hostless or intergalactic SNe in clusters and, in the high-z Universe, would have in situ enriched the CGM and IGM. Further, the hot, young stars formed in outflows will have a large escape fraction of ionising photons which, if they existed in the early Universe, could contribute to the reionisation of the Universe.

Those stars that are formed and are gravitationally bound will instead play a role in the evolution of the spherical component of their host. As the resulting orbits of the stars may be predominantly radial, they will oscillate around the galactic centre. Gravitational interactions with the disk or other non-spherically symmetric potentials would randomise their orbits, contributing significantly to the formation and evolution of the spheroidal component (bulge or halo), and may even contribute to the formation of elliptical galaxies.

What type of galaxies are most likely to host outflows in the local Universe, and what are the likely outflow driving sources in these galaxies? What effects does outflow feedback have on their host? In chapter 6 we explored some of the general properties of outflow-host galaxies in the local Universe. We found that outflows seem to be prevalent in all types of galaxy, with our sample of 168 outflow-host galaxies forming a representative sample of galaxies in the local Universe according to their median BPT classifications. Exploring their locations in the SFR-M_* plane, however, showed a propensity for outflow-host galaxies to fall on or slightly above the MS, with only a few galaxies in the quiescent region. This suggests that outflows are more likely to occur in galaxies with larger amounts of gas in their ISM to hydrodynamically couple to the outflow winds.

We investigated the outflow driving source by determining the median BPT classification of spaxels within 1 PSF of the host galaxy centre. We found that AGN presence in outflow hosts is double that expected from a representative sample of local galaxies, indicating the importance of AGN in driving outflows. However, the majority of driving sources are classified as SF, implying the important of SF-driven outflows in local galaxies. In the SFR-M_* plane, we found that SF-driven outflow hosts tend to be located above the MS, suggesting the importance of starbursts for SF-driven outflows, whilst AGN-driven outflow hosts are located on both the MS and in the quiescent region. This may suggests that AGN are more effective at driving outflows in galaxies on the MS with more ISM gas available.

The propensity for SF-driven outflows to be in galaxies above the MS can have a number of implications. Whilst it is suggestive that starbursts are important for SF-driven outflows, it could also be an indication of outflow feedback inducing SF in the galactic disk.

Our studies of gas densities in the disk of our outflow host galaxies has suggested that outflows do not have a noticeable effect on the density in the disk.

How are the propensity of and properties of outflows influenced by the properties of their host galaxies and their driving source? In chapter 6 we explored the general properties of the outflowing gas in local galactic outflows. We began by investigating the density of the gas entrained within the outflows. We found that the gas in the outflow is, on average, denser than the gas in the disk by a factor of ∼3. Splitting the galaxies by their driving source, we found that SF-driven outflows were less dense, on average, than AGN-driven outflows. As a function of radial distance from the outflow centre, SF-driven outflows were also found to be less dense at all radii than AGN-driven outflows, and in both cases the density decreases with radius.

We found that outflow velocities of SF-driven outflows generally do not to exceed a few hundred km s^{-1} and are lower, on average, than AGN-driven outflows, which can exceed 1000 km s^{-1}.

Our studies of outflow extents suggested that AGN-driven outflows are slightly more extended, on average, than SF-driven outflows, though some of the most extensive outflows are SF-driven, likely due to their driving source being distributed on larger scales. The outflow extent also shows a strong correlation with distance along the MS, suggesting a correlation between both SFR and stellar mass with outflow extent.

For the outflow dynamical timescales we observe that SF-driven outflows have only slightly longer timescales than AGN-driven outflows, likely due to the flickering nature of AGN. The timescales also increase along the MS. This suggests larger galaxies may be able to sustain outflows for longer periods of time.

We also inferred outflow masses and outflow mass rates from both Hα and [OIII] emission. For the Hα outflow masses and rates, we see that AGN-driven outflows are greater in both quantities than SF-driven outflows by a factor of ∼2. In our [OIII] inferred outflow masses and rates we see a similar trend, though in these quentities the AGN-driven outflows are greater than those of SF-driven outflows by a factor ∼10. These results suggests that AGN-driven outflows are more massive and posses higher rates than SF-driven outflows, confirming the important negative feedback role of AGN. The results also show that [OIII] preferentially traces AGN-driven outflows, though this is unsurprising due to the radiation field of AGN increasing [OIII] presence in comparison to lower ionisation lines.

7.2 Future work

Two of the largest limitations of the MaNGA data, however, are the limited angular resolution, which has prevented us from properly resolving the complex internal structures of outflows, and the lack of sensitivity, which has prevented us from properly exploring outflows in low-mass galaxies. We can overcome the issue of limited resolution through future observations

of our outflow-host galaxy sample using IFUs at larger telescopes, such as MUSE at the VLT. Such high angular resolution data will help us to constrain, in detail, the properties of ionised outflows. Furthermore, the increased sensitivity will allow us to detect and classify the weak outflows likely found in low-mass galaxies.

Despite their multi-phase nature, our analyses of outflows throughout this thesis have been limited to the ionised component that, although prominent in many galaxies, may only account for a small fraction of the total outflowing gas in most galaxies. The next logical step would be to investigate the atomic neutral phase of the outflow through mapping the NaID absorption in MaNGA (or potentially more sensitive MUSE) data to infer or constrain the amount of neutral atomic gas in the outflows. For the molecular phase, which is typically the most massive and energetic component of outflows, we would require deep ALMA observations of the CO (or [CI]) transitions of the galaxies in our sample. Finally, we would require deep X-ray observations of our outflows to trace the very hot medium typically found in outflows.

Similar to our investigation into positive feedback in IRAS F23128−5919, as discussed in chapter 4, we would like to follow up the galaxies from the MaNGA sample for which we have detected positive feedback with further observations to confirm the *in situ* star formation, as discussed in chapter 5. Whilst the BPT diagrams provide good evidence for this new channel of SF, supported by the ionisation parameter diagram, these diagnostic are not sufficient alone to unambiguously confirm the presence of stars forming within an outflow. Observations in the NIR, such as those offered by X-shooter, can provide access to the NIR diagnostics that provide a much more clear separation between SF and AGN/shock ionisation. Whilst we have yet to follow up any of the galaxies from the MaNGA sample, we have recently observed a number of outflows with X-shooter for galaxies showing signatures of positive feedback, as determined using BPT diagrams from MUSE observations, which are in the process of being analysed. Furthermore, HST-COS far-UV observations would allow us to directly trace the clear signature of young OB stars formed in the outflows, by finding signatures of blue-shifted photospheric features such as the CIII1176Å line. For IRAS F23128−5919 these such observations have been obtained, and indeed we do detect this feature blue-shifted by 300−1000 km s^{-1}, similar to the outflow velocity (Belfiore et al. *in prep.*).

Finally, we would like to extend our studies to galaxies at higher redshifts ($z \sim 1-3$) where the bulk of both negative and positive feedback due to outflows is expected to take place. At these redshifts, the optical nebular lines used to trace the ionised component of outflows are shifted to the NIR. Efforts are currently being undertaken by various groups using ground based NIR IFU spectrographs, such as SINFONI and KMOS on the VLT. However, the major revolution in this field will soon come with the launch of the JWST, which will enable us to obtain high angular resolution IFU spectra in the NIR with an unprecedented sensitivity.

References

Aalto, S., Garcia-Burillo, S., Muller, S., et al. 2012, A&A, 537, A44

Aalto, S., Martín, S., Costagliola, F., et al. 2015, A&A, 584, A42

Abazajian, K. N., Adelman-McCarthy, J. K., Agüeros, M. A., et al. 2009, ApJS, 182, 543

Abolfathi, B., Aguado, D. S., Aguilar, G., et al. 2018, ApJS, 235, 42

Aguado, D. S., Ahumada, R., Almeida, A., et al. 2019, ApJS, 240, 23

Allen, M. G., Groves, B. A., Dopita, M. A., Sutherland, R. S., & Kewley, L. J. 2008, ApJS, 178, 20

Allington-Smith, J., & Content, R. 1998, PASP, 110, 1216

Anglés-Alcázar, D., Davé, R., Faucher-Giguère, C.-A., Özel, F., & Hopkins, P. F. 2017, MNRAS, 464, 2840

Arribas, S., Colina, L., Bellocchi, E., Maiolino, R., & Villar-Martín, M. 2014, A&A, 568, A14

Bacon, R., Copin, Y., Monnet, G., et al. 2001, MNRAS, 326, 23

Bae, H.-J., & Woo, J.-H. 2014, ApJ, 795, 30

Bae, H.-J., & Woo, J.-H. 2018, ApJ, 853, 185

Baldry, I. K., Glazebrook, K., Brinkmann, J., et al. 2004, ApJ, 600, 681

Baldwin, J. A., Phillips, M. M., & Terlevich, R. 1981, PASP, 93, 5

Balmaverde, B., Marconi, A., Brusa, M., et al. 2016, A&A, 585, A148

Baskin, A., & Laor, A. 2005, MNRAS, 358, 1043

Baug, T., Dewangan, L. K., Ojha, D. K., et al. 2018, ApJ, 852, 119

Belfiore, F., Maiolino, R., Maraston, C., et al. 2016, MNRAS, 461, 3111

Belfiore, F., Westfall, K. B., Schaefer, A., et al. 2019, arXiv:1901.00866

Bell, E. F., Wolf, C., Meisenheimer, K., et al. 2004, ApJ, 608, 752

Bell, E. F., Papovich, C., Wolf, C., et al. 2005, ApJ, 625, 23

Bellocchi, E., Arribas, S., Colina, L., & Miralles-Caballero, D. 2013, A&A, 557, A59

Belokurov, V., Erkal, D., Evans, N. W., Koposov, S. E., & Deason, A. J. 2018, MNRAS, 478, 611

Belokurov, V., Sanders, J. L., Fattahi, A., et al. 2019, arXiv:1909.04679

Benson, A. J., Bower, R. G., Frenk, C. S., et al. 2003, ApJ, 599, 38

Bertola, F., Cinzano, P., Corsini, E. M., Rix, H.-W., & Zeilinger, W. W. 1995, ApJ, 448, L13

Bicknell, G. V., Sutherland, R. S., van Breugel, W. J. M., et al. 2000, ApJ, 540, 678

Bieri, R., Dubois, Y., Silk, J., Mamon, G. A., & Gaibler, V. 2016, MNRAS, 455, 4166

Bîrzan, L., Rafferty, D. A., McNamara, B. R., Wise, M. W., & Nulsen, P. E. J. 2004, ApJ, 607, 800

Bîrzan, L., Rafferty, D. A., Nulsen, P. E. J., et al. 2012, MNRAS, 427, 3468

Bischetti, M., Maiolino, R., Fiore, S. C. F., Piconcelli, E., & Fluetsch, A. 2018, arXiv:1806.00786

Blanton, E. L., Sarazin, C. L., & McNamara, B. R. 2003, ApJ, 585, 227

Blanton, M. R., & Roweis, S. 2007, AJ, 133, 734

Blanton, M. R., Kazin, E., Muna, D., Weaver, B. A., & Price-Whelan, A. 2011, AJ, 142, 31

Blanton, M. R., Bershady, M. A., Abolfathi, B., et al. 2017, AJ, 154, 28

Bluck, A. F. L., et al. 2019b, *submitted*

Boehringer, H., Voges, W., Fabian, A. C., Edge, A. C., & Neumann, D. M. 1993, MNRAS, 264, L25

Bordoloi, R., Lilly, S. J., Hardmeier, E., et al. 2014, ApJ, 794, 130

Borguet, B. C. J., Edmonds, D., Arav, N., Dunn, J., & Kriss, G. A. 2012, ApJ, 751, 107

Boroson, T. 2005, AJ, 130, 381

Bower, R. G., Benson, A. J., Malbon, R., et al. 2006, MNRAS, 370, 645

Bower, R. G., Benson, A. J., & Crain, R. A. 2012, MNRAS, 422, 2816

Bradshaw, E. J., Almaini, O., Hartley, W. G., et al. 2013, MNRAS, 433, 194

Brand, J., Massi, F., Zavagno, A., Deharveng, L., & Lefloch, B. 2011, A&A, 527, A62

Brightman, M., & Nandra, K. 2011, MNRAS, 413, 1206

Brinchmann, J., Charlot, S., White, S. D. M., et al. 2004, MNRAS, 351, 1151

Brusa, M., Bongiorno, A., Cresci, G., et al. 2015, MNRAS, 446, 2394

Brusa, M., Feruglio, C., Cresci, G., et al. 2015, A&A, 578, A11

Brusa, M., Perna, M., Cresci, G., et al. 2016, A&A, 588, A58

Bruzual, G., & Charlot, S. 2003, MNRAS, 344, 1000

Bundy, K., Bershady, M. A., Law, D. R., et al. 2015, ApJ, 798, 7

Calzetti, D., Kinney, A. L., & Storchi-Bergmann, T. 1994, ApJ, 429, 582

Calzetti, D., Armus, L., Bohlin, R. C., et al. 2000, ApJ, 533, 682

Calzetti, D. 2001, PASP, 113, 1449

Cano-Díaz, M., Maiolino, R., Marconi, A., et al. 2012, A&A, 537, L8

Cappellari, M., & Copin, Y. 2003, MNRAS, 342, 345

Cappellari, M., & Emsellem, E. 2004, PASP, 116, 138

Cappellari, M., Emsellem, E., Bacon, R., et al. 2007, MNRAS, 379, 418

Cappellari, M., Emsellem, E., Krajnović, D., et al. 2011, MNRAS, 413, 813
Cappellari, M., Emsellem, E., Krajnović, D., et al. 2011, MNRAS, 416, 1680
Cappellari, M. 2016, ARA&A, 54, 597
Cappellari, M. 2017, MNRAS, 466, 798
Cardelli, J. A., Clayton, G. C., & Mathis, J. S. 1989, ApJ, 345, 245
Carilli, C. L., Perley, R. A., & Harris, D. E. 1994, MNRAS, 270, 173
Carniani, S., Marconi, A., Maiolino, R., et al. 2015, A&A, 580, A102
Carniani, S., Marconi, A., Maiolino, R., et al. 2016, A&A, 591, A28
Cattaneo, A., Faber, S. M., Binney, J., et al. 2009, Nature, 460, 213
Cazzoli, S., Arribas, S., Colina, L., et al. 2014, A&A, 569, A14
Cazzoli, S., Arribas, S., Maiolino, R., & Colina, L. 2016, A&A, 590, A125
Chabrier, G. 2003, PASP, 115, 763
Chang, Y.-Y., van der Wel, A., da Cunha, E., & Rix, H.-W. 2015, ApJS, 219, 8
Chen, Y.-M., Tremonti, C. A., Heckman, T. M., et al. 2010, AJ, 140, 445
Cheung, E., Bundy, K., Cappellari, M., et al. 2016, Nature, 533, 504
Chevalier, R. A. 1977, ARA&A, 15, 175
Chevalier, R. A., & Clegg, A. W. 1985, Nature, 317, 44
Cicone, C., Maiolino, R., Sturm, E., et al. 2014, A&A, 562, A21
Cicone, C., Maiolino, R., Gallerani, S., et al. 2015, A&A, 574, A14
Cicone, C., Brusa, M., Ramos Almeida, C., et al. 2018, Nature Astronomy, 2, 176
Cid Fernandes, R., Stasińska, G., Schlickmann, M. S., et al. 2010, MNRAS, 403, 1036
Cid Fernandes, R., Stasińska, G., Mateus, A., & Vale Asari, N. 2011, MNRAS, 413, 1687
Ciotti, L., Ostriker, J. P., & Proga, D. 2010, ApJ, 717, 708
Colina, L., Piqueras López, J., Arribas, S., et al. 2015, A&A, 578, A48
Combes, F., García-Burillo, S., Casasola, V., et al. 2014, A&A, 565, A97
Concas, A., Popesso, P., Brusa, M., et al. 2017, A&A, 606, A36
Concas, A., Popesso, P., Brusa, M., Mainieri, V., & Thomas, D. 2019, A&A, 622, A188
Corsini, E. M., Pizzella, A., Sarzi, M., et al. 1999, A&A, 342, 671
Costa, T., Sijacki, D., Trenti, M., & Haehnelt, M. G. 2014, MNRAS, 439, 2146
Costa, T., Sijacki, D., & Haehnelt, M. G. 2014, MNRAS, 444, 2355
Costa, T., Sijacki, D., & Haehnelt, M. G. 2015, MNRAS, 448, L30
Cresci, G., & Maiolino, R. 2018, Nature Astronomy, 2, 179
Cresci, G., Mainieri, V., Brusa, M., et al. 2015, ApJ, 799, 82
Cresci, G., Marconi, A., Zibetti, S., et al. 2015, A&A, 582, A63
Crockett, R. M., Shabala, S. S., Kaviraj, S., et al. 2012, MNRAS, 421, 1603
Croft, S., van Breugel, W., de Vries, W., et al. 2006, ApJ, 647, 1040

Croom, S. M., Lawrence, J. S., Bland-Hawthorn, J., et al. 2012, MNRAS, 421, 872
Croton, D. J., Springel, V., White, S. D. M., et al. 2006, MNRAS, 365, 11
Curti, M., Cresci, G., Mannucci, F., et al. 2017, MNRAS, 465, 1384
Daddi, E., Dickinson, M., Morrison, G., et al. 2007, ApJ, 670, 156
Dasyra, K. M., & Combes, F. 2012, A&A, 541, L7
Dasyra, K. M., Combes, F., Oosterloo, T., et al. 2016, A&A, 595, L7
Davé, R., Cen, R., Ostriker, J. P., et al. 2001, ApJ, 552, 473
Dawson, K. S., Schlegel, D. J., Ahn, C. P., et al. 2013, AJ, 145, 10
de Zeeuw, P. T., Bureau, M., Emsellem, E., et al. 2002, MNRAS, 329, 513
Decarli, R., Walter, F., Venemans, B. P., et al. 2018, ApJ, 854, 97
Deharveng, L., Zavagno, A., Samal, M. R., et al. 2015, A&A, 582, A1
Decataldo, D., Ferrara, A., Pallottini, A., Gallerani, S., & Vallini, L. 2017, MNRAS, 471, 4476
Decataldo, D., Pallottini, A., Ferrara, A., Vallini, L., & Gallerani, S. 2019, MNRAS, 487, 3377
Dewangan, L. K., & Ojha, D. K. 2013, MNRAS, 429, 1386
Dewangan, L. K., Ojha, D. K., Anandarao, B. G., Ghosh, S. K., & Chakraborti, S. 2012, ApJ, 756, 151
Dewangan, L. K., Luna, A., Ojha, D. K., et al. 2015, ApJ, 811, 79
Di Matteo, T., Springel, V., & Hernquist, L. 2005, Nature, 433, 604
Díaz, A. I., Castellanos, M., Terlevich, E., & Luisa García-Vargas, M. 2000, MNRAS, 318, 462
Diniz, S. I. F., Pastoriza, M. G., Hernandez-Jimenez, J. A., et al. 2017, MNRAS, 470, 1703
Dopita, M. A., & Sutherland, R. S. 1995, ApJ, 455, 468
Dopita, M. A., Ho, I.-T., Dressel, L. L., et al. 2015, ApJ, 801, 42
Draine, B. T. 2003, A&A, 41, 241
Dressler, A. 1989, Active Galactic Nuclei, 134, 217
Drory, N., MacDonald, N., Bershady, M. A., et al. 2015, AJ, 149, 77
Dugan, Z., Bryan, S., Gaibler, V., Silk, J., & Haas, M. 2014, ApJ, 796, 113
Dunn, R. J. H., & Fabian, A. C. 2008, MNRAS, 385, 757
Duronea, N. U., Cappa, C. E., Bronfman, L., et al. 2017, A&A, 606, A8
Dwarkadas, V. V., Dauphas, N., Meyer, B., Boyajian, P., & Bojazi, M. 2017, ApJ, 851, 147
El-Badry, K., Wetzel, A., Geha, M., et al. 2016, ApJ, 820, 131
Elbaz, D., Jahnke, K., Pantin, E., Le Borgne, D., & Letawe, G. 2009, A&A, 507, 1359
Elvis, M. 2000, ApJ, 545, 63
Emsellem, E., Cappellari, M., Peletier, R. F., et al. 2004, MNRAS, 352, 721
Emsellem, E., Cappellari, M., Krajnović, D., et al. 2007, MNRAS, 379, 401
Fabian, A. C. 1994, ARA&A, 32, 277
Fabian, A. C. 1999, MNRAS, 308, L39

Fabian, A. C. 2012, ARA&A, 50, 455
Fabian, A. C., Wilman, R. J., & Crawford, C. S. 2002, MNRAS, 329, L18
Fabian, A. C., Celotti, A., & Erlund, M. C. 2006, MNRAS, 373, L16
Fabian, A. C., Vasudevan, R. V., & Gandhi, P. 2008, MNRAS, 385, L43
Falcón-Barroso, J., Sánchez-Blázquez, P., Vazdekis, A., et al. 2011, A&A, 532, A95
Faucher-Giguère, C.-A., & Quataert, E. 2012, MNRAS, 425, 605
Feain, I. J., Papadopoulos, P. P., Ekers, R. D., & Middelberg, E. 2007, ApJ, 662, 872
Ferguson, H. C., Tanvir, N. R., & von Hippel, T. 1998, Nature, 391, 461
Ferland, G. J., & Netzer, H. 1983, ApJ, 264, 105
Ferrarese, L., & Merritt, D. 2000, ApJ, 539, L9
Feruglio, C., Maiolino, R., Piconcelli, E., et al. 2010, A&A, 518, L155
Figueira, M., Zavagno, A., Deharveng, L., et al. 2017, A&A, 600, A93
Finn, C. W., Morris, S. L., Crighton, N. H. M., et al. 2014, MNRAS, 440, 3317
Fiore, F., Feruglio, C., Shankar, F., et al. 2017, A&A, 601, A143
Fischer, J., Sturm, E., González-Alfonso, E., et al. 2010, A&A, 518, L41
Fitzpatrick, E. L. 1999, PASP, 111, 63
Fluetsch, A., Maiolino, R., Carniani, S., et al. 2019, MNRAS, 483, 4586
Fluetsch, A., et al. 2019, *in prep.*
Förster Schreiber, N. M., Übler, H., Davies, R. L., et al. 2019, ApJ, 875, 21
Fu, H., & Stockton, A. 2009, ApJ, 690, 953
Fumagalli, M., Fossati, M., Hau, G. K. T., et al. 2014, MNRAS, 445, 4335
Gabor, J. M., Davé, R., Oppenheimer, B. D., & Finlator, K. 2011, MNRAS, 417, 2676
Gaibler, V., Khochfar, S., Krause, M., & Silk, J. 2012, MNRAS, 425, 438
Gal-Yam, A., Maoz, D., Guhathakurta, P., & Filippenko, A. V. 2003, AJ, 125, 1087
Gallagher, R., Maiolino, R., Belfiore, F., et al. 2019, MNRAS, 485, 3409
García-Burillo, S., Combes, F., Usero, A., et al. 2015, A&A, 580, A35
Genzel, R., Förster Schreiber, N. M., Rosario, D., et al. 2014, AJp, 796, 7
Girardi, L., Bressan, A., Bertelli, G., et al. 2000, A&AS, 141, 371
González Delgado, R. M., Cerviño, M., Martins, L. P., Leitherer, C., & Hauschildt, P. H. 2005, MNRAS, 357, 945
González-Alfonso, E., Fischer, J., Spoon, H. W. W., et al. 2017, ApJ, 836, 11
Graham, M. L., Sand, D. J., Zaritsky, D., & Pritchet, C. J. 2015, ApJ, 807, 83
Granato, G. L., De Zotti, G., Silva, L., Bressan, A., & Danese, L. 2004, ApJ, 600, 580
Guérou, A., Emsellem, E., McDermid, R. M., et al. 2015, ApJ, 804, 70
Guérou, A., Emsellem, E., Krajnović, D., et al. 2016, A&A, 591, A143
Gunn, J. E., Siegmund, W. A., Mannery, E. J., et al. 2006, AJ, 131, 2332

Gupta, R. R., Kuhlmann, S., Kovacs, E., et al. 2016, AJ, 152, 154

Gutiérrez, L., & Beckman, J. E. 2010, ApJ, 710, L44

Häring, N., & Rix, H.-W. 2004, ApJ, 604, L89

Harrison, C. M., Alexander, D. M., Swinbank, A. M., et al. 2012, MNRAS, 426, 1073

Harrison, C. M., Alexander, D. M., Mullaney, J. R., & Swinbank, A. M. 2014, MNRAS, 441, 3306

Harrison, C. M., Alexander, D. M., Mullaney, J. R., et al. 2016, MNRAS, 456, 1195

Harrison, C. M., Costa, T., Tadhunter, C. N., et al. 2018, Nature Astronomy, 2, 198

Heckman, T. M. 1980, A&A, 87, 152

Heckman, T. M., & Best, P. N. 2014, ARA&A, 52, 589

Heckman, T. M., & Borthakur, S. 2016, ApJ, 822, 9

Heckman, T. M., & Thompson, T. A. 2017, arXiv:1701.09062

Heckman, T. M., Miley, G. K., & Green, R. F. 1984, ApJ, 281, 525

Heckman, T. M., Armus, L., & Miley, G. K. 1987, AJ, 93, 276

Heckman, T. M., Armus, L., & Miley, G. K. 1990, ApJS, 74, 833

Heckman, T. M., Alexandroff, R. M., Borthakur, S., Overzier, R., & Leitherer, C. 2015, ApJ, 809, 147

Hinshaw, G., Larson, D., Komatsu, E., et al. 2013, ApJS, 208, 19

Hlavacek-Larrondo, J., Fabian, A. C., Edge, A. C., et al. 2012, MNRAS, 421, 1360

Hlavacek-Larrondo, J., Allen, S. W., Taylor, G. B., et al. 2013, ApJ, 777, 163

Ho, L. C., Filippenko, A. V., & Sargent, W. L. W. 1993, ApJ, 417, 63

Ho, I.-T., Medling, A. M., Groves, B., et al. 2016, Ap&SS, 361, 280

Hopkins, P. F., Hernquist, L., Cox, T. J., et al. 2006, ApJS, 163, 1

Hubble, E. P. 1926, ApJ, 64,

Hubble, E. P. 1936, Realm of the Nebulae, by E.P. Hubble. New Haven: Yale University Press, 1936. ISBN 9780300025002,

Husemann, B., Jahnke, K., Sánchez, S. F., et al. 2013, A&A, 549, A87

Ishibashi, W., & Fabian, A. C. 2012, MNRAS, 427, 2998

Ishibashi, W., & Fabian, A. C. 2014, MNRAS, 441, 1474

Ishibashi, W., & Fabian, A. C. 2015, MNRAS, 451, 93

Ishibashi, W., & Fabian, A. C. 2017, MNRAS, 472, 2768

Ishibashi, W., Fabian, A. C., & Canning, R. E. A. 2013, MNRAS, 431, 2350

Ishibashi, W., Fabian, A. C., & Maiolino, R. 2018, MNRAS, 476, 512

Janssen, A. W., Christopher, N., Sturm, E., et al. 2016, ApJ, 822, 43

Kakkad, D., Groves, B., Dopita, M., et al. 2018, A&A, 618, A6

Kauffmann, G., Heckman, T. M., White, S. D. M., et al. 2003b, MNRAS, 341, 33

Kauffmann, G., Heckman, T. M., Tremonti, C., et al. 2003c, MNRAS, 346, 1055

Kauffmann, G., & Haehnelt, M. 2000, MNRAS, 311, 576

Kennicutt, R. C., Jr. 1998, ARA&A, 36, 189

Kennicutt, R. C., & Evans, N. J. 2012, ARA&A, 50, 531

Kewley, L. J., Dopita, M. A., Sutherland, R. S., Heisler, C. A., & Trevena, J. 2001, ApJ, 556, 121

Kewley, L. J., Groves, B., Kauffmann, G., & Heckman, T. 2006, MNRAS, 372, 961

King, A. 2003, ApJ, 596, L27

King, A. 2005, ApJ, 635, L121

King, A. R. 2010, MNRAS, 402, 1516

King, A., & Pounds, K. 2015, ARA&A, 53, 115

King, A. R., Zubovas, K., & Power, C. 2011, MNRAS, 415, L6

Kinney, A. L., Schmitt, H. R., Clarke, C. J., et al. 2000, ApJ, 537, 152

Komossa, S., Xu, D., Zhou, H., Storchi-Bergmann, T., & Binette, L. 2008, ApJ, 680, 926

Kormendy, J. 1993, The Nearest Active Galaxies, 197

Kormendy, J., & Bender, R. 2009, ApJ, 691, L142

Kormendy, J., & Bender, R. 2011, Nature, 469, 377

Kormendy, J., & Bender, R. 2012, ApJS, 198, 2

Kormendy, J., & Gebhardt, K. 2001, 20th Texas Symposium on relativistic astrophysics, 586, 363

Kormendy, J., & Ho, L. C. 2013, ARA&A, 51, 511

Kroupa, P. 2001, MNRAS, 322, 231

Kroupa, P., Tout, C. A., & Gilmore, G. 1993, MNRAS, 262, 545

Kumari, N., Maiolino, R., Belfiore, F., & Curti, M. 2019, MNRAS, 485, 367

Lacy, M., Croft, S., Fragile, C., Wood, S., & Nyland, K. 2017, ApJ, 838, 146

Ladeyschikov, D. A., Sobolev, A. M., Parfenov, S. Y., Alexeeva, S. A., & Bieging, J. H. 2015, MNRAS, 452, 2306

LaMassa, S. M., Glikman, E., Brusa, M., et al. 2017, ApJ, 847, 100

Lapi, A., Cavaliere, A., & Menci, N. 2005, ApJ, 619, 60

Law, D. R., Yan, R., Bershady, M. A., et al. 2015, AJ, 150, 19

Law, D. R., Cherinka, B., Yan, R., et al. 2016, AJ, 152, 83

Lehnert, M. D., & Heckman, T. M. 1996, ApJ, 462, 651

Leitherer, C., Robert, C., & Drissen, L. 1992, ApJ, 401, 596

Leslie, S. K., Rich, J. A., Kewley, L. J., & Dopita, M. A. 2014, MNRAS, 444, 1842

Leung, G. C. K., Coil, A. L., Azadi, M., et al. 2017, ApJ, 849, 48

Lilly, S. J., Carollo, C. M., Pipino, A., Renzini, A., & Peng, Y. 2013, ApJ, 772, 119

Lilly, S. J., Le Fevre, O., Hammer, F., & Crampton, D. 1996, ApJ, 460, L1

Lim, B., Sung, H., Bessell, M. S., et al. 2018, MNRAS, 477, 1993

Liu, G., Zakamska, N. L., Greene, J. E., Nesvadba, N. P. H., & Liu, X. 2013, MNRAS, 436, 2576

Mac Low, M.-M., & Ferrara, A. 1999, ApJ, 513, 142

Madau, P., & Shull, J. M. 1996, ApJ, 457, 551

Magorrian, J., Tremaine, S., Richstone, D., et al. 1998, AJ, 115, 2285

Maiolino, R., Gallerani, S., Neri, R., et al. 2012, MNRAS, 425, L66

Maiolino, R., Russell, H. R., Fabian, A. C., et al. 2017, Nature, 544, 202

Marconi, A., & Hunt, L. K. 2003, ApJ, 589, L21

Marconi, A., Risaliti, G., Gilli, R., et al. 2004, MNRAS, 351, 169

Martinsson, T. P. K., Verheijen, M. A. W., Westfall, K. B., et al. 2013, A&A, 557, A131

Masters, D., Faisst, A., & Capak, P. 2016, ApJ, 828, 18

Mathis, J. S. 1990, ARA&A, 28, 37

McDermid, R. M., Alatalo, K., Blitz, L., et al. 2015, MNRAS, 448, 3484

McNamara, B. R., & Nulsen, P. E. J. 2007, ARA&A, 45, 117

McNamara, B. R., & Nulsen, P. E. J. 2012, New Journal of Physics, 14, 055023

Menci, N., Fontana, A., Giallongo, E., Grazian, A., & Salimbeni, S. 2006, ApJ, 647, 753

Menci, N., Fiore, F., Puccetti, S., & Cavaliere, A. 2008, ApJ, 686, 219

Merritt, D., & Ferrarese, L. 2001, MNRAS, 320, L30

Mingozzi, M., Cresci, G., Venturi, G., et al. 2019, A&A, 622, A146

Miyamoto, M., & Nagai, R. 1975, PASJ, 27, 533

Mo, H., van den Bosch, F. C., & White, S. 2010, Galaxy Formation and Evolution, by Houjun Mo , Frank van den Bosch , Simon White, Cambridge, UK: Cambridge University Press, 2010,

Modigliani, A., Goldoni, P., Royer, F., et al. 2010, Proc. SPIE, 7737, 773728

Molnár, D. C., Sargent, M. T., Elbaz, D., Papadopoulos, P. P., & Silk, J. 2017, MNRAS, 467, 586

Monreal-Ibero, A., Arribas, S., Colina, L., et al. 2010, A&A, 517, A28

Morganti, R., Oosterloo, T. A., Tadhunter, C. N., van Moorsel, G., & Emonts, B. 2005, A&A, 439, 521

Morganti, R., Oosterloo, T., Oonk, J. B. R., Frieswijk, W., & Tadhunter, C. 2015, A&A, 580, A1

Morganti, R., Veilleux, S., Oosterloo, T., Teng, S. H., & Rupke, D. 2016, A&A, 593, A30

Morselli, L., Popesso, P., Erfanianfar, G., & Concas, A. 2017, A&A, 597, A97

Mukherjee, D., Bicknell, G. V., Wagner, A. Y., Sutherland, R. S., & Silk, J. 2018, MNRAS,

479, 5544

Mullaney, J. R., Alexander, D. M., Fine, S., et al. 2013, MNRAS, 433, 622

Murray, N., Quataert, E., & Thompson, T. A. 2005, ApJ, 618, 569

Myeong, G. C., Evans, N. W., Belokurov, V., Sanders, J. L., & Koposov, S. E. 2018, ApJ, 856, L26

Nagao, T., Maiolino, R., & Marconi, A. 2006, A&A, 459, 85

Natta, A., & Panagia, N. 1984, ApJ, 287, 228

Nayakshin, S. 2014, MNRAS, 437, 2404

Nayakshin, S., & Zubovas, K. 2012, MNRAS, 427, 372

Nesvadba, N. P. H., Lehnert, M. D., De Breuck, C., Gilbert, A. M., & van Breugel, W. 2008, A&A, 491, 407

Nesvadba, N. P. H., De Breuck, C., Lehnert, M. D., et al. 2011, A&A, 525, A43

Newman, S. F., Genzel, R., Förster-Schreiber, N. M., et al. 2012, ApJ, 761, 43

Noeske, K. G., Weiner, B. J., Faber, S. M., et al. 2007, ApJ, 660, L43

Oh, K., Sarzi, M., Schawinski, K., & Yi, S. K. 2011, ApJS, 195, 13

Oliva, E., Marconi, A., Maiolino, R., et al. 2001, A&A, 369, L5

Osterbrock, D. E., & Ferland, G. J. 2006, Astrophysics of gaseous nebulae and active galactic nuclei, 2nd. ed. by D.E. Osterbrock and G.J. Ferland. Sausalito, CA: University Science Books, 2006,

Papaderos, P., Gomes, J. M., Vílchez, J. M., et al. 2013, A&A, 555, L1

Papastergis, E., Cattaneo, A., Huang, S., Giovanelli, R., & Haynes, M. P. 2012, ApJ, 759, 138

Peng, C. Y. 2007, ApJ, 671, 1098

Peng, Y.-j., Lilly, S. J., Kovač, K., et al. 2010, ApJ, 721, 193

Peng, Y., Maiolino, R., & Cochrane, R. 2015, Nature, 521, 192

Pereira-Santaella, M., Colina, L., García-Burillo, S., et al. 2016, A&A, 594, A81

Pérez-Montero, E., & Contini, T. 2009, MNRAS, 398, 949

Perna, M., Brusa, M., Cresci, G., et al. 2015, A&A, 574, A82

Perna, M., Lanzuisi, G., Brusa, M., Cresci, G., & Mignoli, M. 2017, A&A, 606, A96

Peterson, J. R., & Fabian, A. C. 2006, Phys. Rep., 427, 1

Piqueras López, J., Colina, L., Arribas, S., Alonso-Herrero, A., & Bedregal, A. G. 2012, A&A, 546, A64

Planck Collaboration, Aghanim, N., Akrami, Y., et al. 2018, arXiv:1807.06209

Poggianti, B. M., Fasano, G., Omizzolo, A., et al. 2016, AJ, 151, 78

Pradhan, A. K., Montenegro, M., Nahar, S. N., & Eissner, W. 2006, MNRAS, 366, L6

Privon, G. C., Aalto, S., Falstad, N., et al. 2017, ApJ, 835, 213

Protassov, R., van Dyk, D. A., Connors, A., Kashyap, V. L., & Siemiginowska, A. 2002, ApJ,

571, 545
Proxauf, B., Öttl, S., & Kimeswenger, S. 2014, A&A, 561, A10
Rafferty, D. A., McNamara, B. R., Nulsen, P. E. J., & Wise, M. W. 2006, ApJ, 652, 216
Rafferty, D. A., McNamara, B. R., & Nulsen, P. E. J. 2008, ApJ, 687, 899
Read, J. I., & Trentham, N. 2005, Philosophical Transactions of the Royal Society of London Series A, 363,
Rees, M. J. 1989, MNRAS, 239, 1P
Renzini, A., & Peng, Y.-j. 2015, ApJ, 801, L29
Richings, A. J., & Faucher-Giguère, C.-A. 2018, MNRAS, 474, 3673
Rizza, E., Loken, C., Bliton, M., et al. 2000, AJ, 119, 21
Rodighiero, G., Daddi, E., Baronchelli, I., et al. 2011, ApJ, 739, L40
Rodríguez-Zaurín, J., Arribas, S., Monreal-Ibero, A., et al. 2011, A&A, 527, A60
Rodríguez del Pino, B., Arribas, S., Piqueras López, J., Villar-Martín, M., and Colina, L. 2019, MNRAS, 486, 344.
Roth, N., Kasen, D., Hopkins, P. F., & Quataert, E. 2012, ApJ, 759, 36
Rubin, K. H. R., Prochaska, J. X., Koo, D. C., et al. 2014, ApJ, 794, 156
Rupke, D. S. N., & Veilleux, S. 2011, ApJ, 729, L27
Rupke, D. S. N., & Veilleux, S. 2013, ApJ, 768, 75
Rupke, D. S. N., & Veilleux, S. 2015, ApJ, 801, 126
Rupke, D. S., Veilleux, S., & Sanders, D. B. 2005, ApJS, 160, 115
Rupke, D. S., Veilleux, S., & Sanders, D. B. 2005, ApJ, 632, 751
Sakamoto, K., Aalto, S., Wilner, D. J., et al. 2009, ApJ, 700, L104
Sakamoto, K., Aalto, S., Combes, F., Evans, A., & Peck, A. 2014, ApJ, 797, 90
Salim, S., Rich, R. M., Charlot, S., et al. 2007, ApJS, 173, 267
Salomé, Q., Salomé, P., & Combes, F. 2015, A&A, 574, A34
Salpeter, E. E. 1955, ApJ, 121, 161
Sánchez-Blázquez, P., Peletier, R. F., Jiménez-Vicente, J., et al. 2006, MNRAS, 371, 703
Sánchez, S. F., Kennicutt, R. C., Gil de Paz, A., et al. 2012a, A&A, 538, A8
Sánchez, S. F., Rosales-Ortega, F. F., Marino, R. A., et al. 2012b, A&A, 546, A2
Santoro, F., Oonk, J. B. R., Morganti, R., & Oosterloo, T. 2015, A&A, 574, A89
Santoro, F., Oonk, J. B. R., Morganti, R., Oosterloo, T. A., & Tadhunter, C. 2016, A&A, 590, A37
Sarzi, M., Rix, H.-W., Shields, J. C., et al. 2005, ApJ, 628, 169
Sarzi, M., Falcón-Barroso, J., Davies, R. L., et al. 2006, MNRAS, 366, 1151
Sarzi, M., Shields, J. C., Schawinski, K., et al. 2010, MNRAS, 402, 2187
Sarzi, M., Kaviraj, S., Nedelchev, B., et al. 2016, MNRAS, 456, L25

Scannapieco, C., Wadepuhl, M., Parry, O. H., et al. 2012, MNRAS, 423, 1726
Schechter, P. 1976, ApJ, 203, 297
Schulz, R., Morganti, R., Nyland, K., et al. 2018, A&A, 617, A38
Shankar, F., Bernardi, M., Sheth, R. K., et al. 2016, MNRAS, 460, 3119
Shepard 1968, Proc. 1968 ACM National Conference, 1, 517
Shapley, A. E., Steidel, C. C., Pettini, M., & Adelberger, K. L. 2003, ApJ, 588, 65
Shapley, A. E., Reddy, N. A., Kriek, M., et al. 2015, ApJ, 801, 88
Schawinski, K., Urry, C. M., Simmons, B. D., et al. 2014, MNRAS, 440, 889
Shetty, S., & Cappellari, M. 2015, MNRAS, 454, 1332
Silk, J. 2013, ApJ, 772, 112
Silk, J., & Mamon, G. A. 2012, RAA, 12, 917
Silk, J., & Norman, C. 2009, ApJ, 700, 262
Silk, J., & Rees, M. J. 1998, A&A, 331, L1
Speagle, J. S., Steinhardt, C. L., Capak, P. L., & Silverman, J. D. 2014, ApJS, 214, 15
Spoon, H. W. W., Farrah, D., Lebouteiller, V., et al. 2013, ApJ, 775, 127
Springel, V., & Hernquist, L. 2003, MNRAS, 339, 289
Springel, V., Di Matteo, T., & Hernquist, L. 2005, MNRAS, 361, 776
Stasińska, G., Cid Fernandes, R., Mateus, A., Sodré, L., & Asari, N. V. 2006, MNRAS, 371, 972
Strateva, I., Ivezić, Ž., Knapp, G. R., et al. 2001, AJ, 122, 1861
Strauss, M. A., Weinberg, D. H., Lupton, R. H., et al. 2002, AJ, 124, 1810
Strom, A. L., Steidel, C. C., Rudie, G. C., et al. 2017, ApJ, 836, 164
Strom, A. L., Steidel, C. C., Rudie, G. C., Trainor, R. F., & Pettini, M. 2018, ApJ, 868, 117
Sturm, E., González-Alfonso, E., Veilleux, S., et al. 2011, ApJ, 733, L16
Swinbank, A. M., Webb, T. M., Richard, J., et al. 2009, MNRAS, 400, 1121
Talia, M., Mignoli, M., Cimatti, A., et al. 2012, A&A, 539, A61
Talia, M., Brusa, M., Cimatti, A., et al. 2017, MNRAS, 471, 4527
Thompson, M. A., Urquhart, J. S., Moore, T. J. T., & Morgan, L. K. 2012, MNRAS, 421, 408
Thompson, T. A., Fabian, A. C., Quataert, E., & Murray, N. 2015, MNRAS, 449, 147
Thornton, K., Gaudlitz, M., Janka, H.-T., & Steinmetz, M. 1998, ApJ, 500, 95
Toba, Y., Bae, H.-J., Nagao, T., et al. 2017, ApJ, 850, 140
Tombesi, F., Cappi, M., Reeves, J. N., et al. 2010, A&A, 521, A57
Tremonti, C. A., Heckman, T. M., Kauffmann, G., et al. 2004, ApJ, 613, 898
Ueda, Y., Hashimoto, Y., Ichikawa, K., et al. 2015, ApJ, 815, 1
van der Marel, R. P., & Franx, M. 1993, ApJ, 407, 525
van der Marel, R. P., Rix, H. W., Carter, D., et al. 1994, MNRAS, 268, 521

Vaona, L., Ciroi, S., Di Mille, F., et al. 2012, MNRAS, 427, 1266
Vazdekis, A., Sánchez-Blázquez, P., Falcón-Barroso, J., et al. 2010, MNRAS, 404, 1639
Veilleux, S., & Osterbrock, D. E. 1987, ApJS, 63, 295
Veilleux, S. 1991, ApJS, 75, 383
Veilleux, S., Cecil, G., & Bland-Hawthorn, J. 2005, ARA&A, 43, 769
Veilleux, S., Meléndez, M., Sturm, E., et al. 2013, ApJ, 776, 27
Venturi, G., Nardini, E., Marconi, A., et al. 2018, A&A, 619, A74
Venturi, G., et al. 2019, *in prep.*
Villar Martín, M., Emonts, B., Humphrey, A., Cabrera Lavers, A., & Binette, L. 2014, MNRAS, 440, 3202
Villar Martín, M., Bellocchi, E., Stern, J., et al. 2015, MNRAS, 454, 439
Wake, D. A., Bundy, K., Diamond-Stanic, A. M., et al. 2017, AJ, 154, 86
Walch, S., & Naab, T. 2015, MNRAS, 451, 2757
Walter, F., Bolatto, A. D., Leroy, A. K., et al. 2017, ApJ, 835, 265
Wang, X., & Loeb, A. 2018, New A, 61, 95
Wang, R., Wagg, J., Carilli, C. L., et al. 2013, ApJ, 773, 44
Weedman, D. W. 1970, ApJ, 159, 405
Weller, J., Ostriker, J. P., Bode, P., & Shaw, L. 2005, MNRAS, 364, 823
Westfall, K. B., Cappellari, M., Bershady, M. A., et al. 2019, arXiv:1901.00856
Woo, J.-H., Bae, H.-J., Son, D., & Karouzos, M. 2016, ApJ, 817, 108
Wright, E. L., Eisenhardt, P. R. M., Mainzer, A. K., et al. 2010, AJ, 140, 1868
Wylezalek, D., Zakamska, N. L., Greene, J. E., et al. 2018, MNRAS, 474, 1499
Yan, R., Tremonti, C., Bershady, M. A., et al. 2016, AJ, 151, 8
Yan, R., Bundy, K., Law, D. R., et al. 2016, AJ, 152, 197
York, D. G., Adelman, J., Anderson, J. E., Jr., et al. 2000, AJ, 120, 1579
Young, L. M., Bureau, M., & Cappellari, M. 2008, ApJ, 676, 317
Zakamska, N. L., & Greene, J. E. 2014, MNRAS, 442, 784
Zakamska, N. L., Hamann, F., Pâris, I., et al. 2016, MNRAS, 459, 3144
Zavagno, A., Russeil, D., Motte, F., et al. 2010, A&A, 518, L81
Zavagno, A., Anderson, L. D., Russeil, D., et al. 2010, A&A, 518, L101
Zhang, D., & Thompson, T. A. 2012, MNRAS, 424, 1170
Zhang, K., Dong, X.-B., Wang, T.-G., & Gaskell, C. M. 2011, ApJ, 737, 71
Zhang, K., Yan, R., Bundy, K., et al. 2017, MNRAS, 466, 3217
Zubovas, K., & King, A. 2012, ApJ, 745, L34
Zubovas, K., & King, A. R. 2014, MNRAS, 439, 400
Zubovas, K., Nayakshin, S., Sazonov, S., & Sunyaev, R. 2013a, MNRAS, 431, 793

Zubovas, K., Nayakshin, S., King, A., & Wilkinson, M. 2013b, MNRAS, 433, 3079
Zubovas, K., Wynn, G. A., & Gualandris, A. 2013c, ApJ, 771, 118